Virtual Inertia Control of Microgrids Using Deep Reinforcement Learning Methods

Vjatšeslav Škiparev

Virtual Inertia Control of Microgrids Using Deep Reinforcement Learning Methods

VJATŠESLAV ŠKIPAREV

TAL
TECH
PRESS

TALLINN UNIVERSITY OF TECHNOLOGY
School of Information Technologies
Department of Software Science

The dissertation was accepted for the defence of the degree of Doctor of Philosophy on 7 August 2023

Supervisor: Prof. Juri Belikov,
Department of Software Science, School of Information Technology
Tallinn University of Technology
Tallinn, Estonia

Co-supervisor: Prof. Eduard Petlenkov,
Department of Computer Systems, School of Information Technology
Tallinn University of Technology
Tallinn, Estonia

Opponents: Prof. Gaber Magdy,
Aswan University
Aswan, Egypt

Prof. João Martins,
Nova University of Lisbon
Caparica, Portugal

Defence of the thesis: 14 September 2023, Tallinn

Declaration:
I hereby declare that this doctoral thesis, my original investigation and achievement, submitted for the doctoral degree at Tallinn University of Technology, has not been submitted for any academic degree elsewhere.

Vjatšeslav Škiparev

_____
signature

Mikrovõrkude virtuaalse inertsi juhtimine sügava stiimulõppe meetoditega

VJATŠESLAV ŠKIPAREV

TAL
TECH
KIRJASTUS

Contents

List of Publications

Thesis based on the following publications:

I K. Nosrati, V. Skiparev, A. Tepljakov, E. Petlenkov, Y. Levron, and J. Belikov. 9 - application of neural network based variable fractional order pid controllers for load frequency control in isolated microgrids. In D. K. Mishra, L. Li, J. Zhang, and M. J. Hossain, editors, *Power System Frequency Control*, pages 203–216. Academic Press, 2023

II V. Skiparev, K. Nosrati, J. Belikov, A. Tepljakov, and E. Petlenkov. An enhanced NN-based load frequency control design of MGs: A fractional order modeling method. In *2023 IEEE International Conference on Compatibility, Power Electronics and Power Engineering (CPE-POWERENG)*, Tallinn, Estonia, 2023

III K. Nosrati, V. Skiparev, A. Tepljakov, E. Petlenkov, J. Belikov, and Y. Levron. Constrained intelligent frequency control in an AC microgrid: An online reinforcement learning based PID tuning approach. In *2023 IEEE Power and Energy Society General Meeting*, pages 1–5, 2023

IV V. Skiparev, K. Nosrati, A. Tepljakov, E. Petlenkov, Y. Levron, J. Belikov, and J. M. Guerrero. Virtual inertia control of an isolated microgrid using NN-VFOPID controller: A self-tuning approach. *IEEE Transactions on Sustainable Energy*, 2023

V K. Nosrati, V. Skiparev, A. Tepljakov, E. Petlenkov, Y. Levron, and J. Belikov. Coordinated PI-based frequency deviation control of isolated hybrid microgrid: An online multi-agent tuning approach via reinforcement learning. In *2022 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*, pages 1–5, 2022

VI V. Skiparev, J. Belikov, E. Petlenkov, and Y. Levron. Reinforcement learning based MIMO controller for virtual inertia control in isolated microgrids. In *IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*, Novi Sad, Serbia, 2022

VII V. Skiparev, R. Machlev, N. R. Chowdhury, Y. Levron, E. Petlenkov, , and J. Belikov. Virtual inertia control methods in islanded microgrids. *Energies*, 14(6):1562, 2021

VIII V. Skiparev, J. Belikov, and E. Petlenkov. MIMO reinforcement learning based approach for frequency support in microgrids with high renewable energy penetration. In *IEEE PES General Meeting*, Washington DC, USA, 2021

IX V. Skiparev, J. Belikov, and E. Petlenkov. Reinforcement learning based approach for virtual inertia control in microgrids with renewable energy sources. In *IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe)*, The Hague, NL, 2020

Approbation

I presented the results of the thesis at the following conferences:

1. **V. Skiparev**, J. Belikov, and E. Petlenkov. 'Reinforcement learning based approach for virtual inertia control in microgrids with renewable energy sources', IEEE ISGT Europe 2020. 8–12 October 2020, Hague, Netherlands

2. **V. Skiparev**, J. Belikov, and E. Petlenkov. 'MIMO reinforcement learning based approach for frequency support in microgrids with high renewable energy penetration', IEEE PES General Meeting 2021, 25-29 July 2021, Washington D.C., USA

3. **V. Skiparev**, J. Belikov, and E. Petlenkov, Y. Levron. 'Reinforcement learning based MIMO controller for virtual inertia control in isolated microgrids', IEEE ISGT Europe 2022, 10-12 October 2022, Novi Sad, Serbia

4. **V. Skiparev**, K. Nosrati, A. Tepljakov, E. Petlenkov, J. Belikov. 'An Enhanced NN-based Load Frequency Control Design of MGs: A Fractional order Modeling Metho', IEEE CPE-POWERENG 2023, 14-16 June 2023, Tallinn, Estonia

5. **K. Nosrati**, V. Skiparev, and A. Teplajkov, and E. Petlenkov and Y. Levron, J. Belikov. 'Constrained Intelligent Frequency Control in an AC Microgrid: An Online Reinforcement Learning Based PID Tuning Approach', IEEE PES General Meeting 2023, 16-20 July 2023, Orlando, USA

Author's Contributions to the Publications

- In [84], I was the first author and proposed the reinforcement learning based controller for virtual inertia control applied for frequency support in the model of the isolated microgrid, where the training was performed by the DDPG policy.

- In [85], I was the first author and proposed reinforcement learning based controller, which provides simultaneous virtual inertia and renewable energy flow control in the isolated microgrid.

- In [87], I was the first author and provided the state of the art on various virtual inertia control methods (including CDM, PID, FLC, MPC, ANN and H_∞) supported by the comparative analysis of the revised control methods.

- In [86], I was the first author and combined the energy storage scheme from [47] with the control method from [85], and developed a new SDDPG training method for virtual inertia emulation in isolated microgrid with high penetration of renewable energy sources.

- In [67], I was the first author and developed an artificial neural network based tuner and combined it with the variable FOPID controller applied for the virtual inertia control scheme proposed in [36] for isolated microgrid with high penetration of renewable energy sources.

- In [89], I was the second author and combined an artificial neural network based tuner with two online tuned PI controllers to provide coordinated control of fuel-cell and double-layer capacitor in isolated microgrid trained by multi-agent reinforcement learning architecture.

- In [66], I was the second author and applied stability analysis of an NN-PI(D) controller to define its stability limitations and guarantee the robustness of the controller.

- In [88], I was the first author and applied multi-agent reinforcement learning to train multiple ANN-based controllers applied for frequency control in hybrid microgrid.

- In [68], I was the second author and applied multi-agent reinforcement learning to train multiple NN-VFOPID controllers applied for frequency control in hybrid microgrid.

Abbreviations

ANN	Artificial Neural Network
ACE	Area control error
A3C	Asynchronous advantage actor-critic
CDM	Coefficient diagram method
RL	Reinforcement learning
DRL	Deep RL
DQL	Deep Q-learning
DLC	Double-layered capacitor
VIC	Virtual inertia controller
VSG	Virtual synchronous generator
MG	Microgrid
ML	Machine Learning
MIMO	Multi Input Multi Output
HGM	Hybrid microgrid
HPF	High-Pass Filter
BSS	Battery storage system
ESS	Energy storage system
DDPG	Deep Deterministic Policy Gradient
SDDPG	Simplified DDPG
SARSA	Step Action Reward Step Action
RES	Renewable energy system
VRE	Variable renewable energy
FOPID	Fractional Order PID
VFOPID	Variable FOPID
RoCoF	Rate of Change of Frequency
PID	Proportional integral differential
PSO	Particle swarm optimisation
PLL	Phase lock-loop
FC	Fuel cell
FLC	Fuzzy-logic controller
MPC	Model predictive control
SRL	Stochastic reinforcement learning
SAC	Soft Actor Critic
PCL	Primary control loop
SISO	Single Input Single Output
SCL	Secondary control loop
LFC	Load frequency control
WTG	Wind turbine generator

Symbols

a_t Action
s_t State
r_t Computed reward
$\mu(s_t|\theta_\mu)$ Actor network
$Q(s_t,a_t|\theta_Q)$ Critic network
$Q(s_{t+1},a_t)$ Best predicted reward by critic network
L_μ Loss function of actor network
L_Q Loss function of critic network
K_{VI} Virtual inertia coefficient
Δf Frequency deviation
$RoCoF$ Rate of Change of Frequency
K_i Integral coefficient of secondary control loop
$1/R$ Droop control
θ_μ Matrix of actor weights
θ_Q Matrix of critic weights
H System inertia
D System damping
R_{VI} Virtual droop
D_{VI} Virtual damping
$\Delta\alpha$ ESS positive feedback control signal
$\Delta\beta$ ESS negative feedback control signal
$\Delta\gamma$ ESS charge control signal
ΔP_L Load power changes
ΔP_M Mechanical power deviation
ΔP_E Electrical power deviation
ΔP_{VI} Virtual inertia power injection
ΔP_{PCL} Control changes of PCL
ΔP_{SCL} ACE action changes of SCL
ΔP_{TPP} Thermal power plant power
ΔP_{PV} Photo-voltaic power
ΔP_{WTG} Wind turbine generator
ΔP_{Wind} Wind power deviation
ΔP_{Solar} Solar power deviation
ΔP_{FC} Fuel cell power deviation
ΔP_{DLC} Double layer capacitor power deviation
ΔP_{RES} Renewable energy sources power deviation
ΔP_g Governor dead band limits
ΔP_{RL} Residential load
ΔP_{IL} Industrial load
T_G Governor time constant
T_t Turbine time constant
T_{WTG} WTG time constant
T_{VI} Virtual inertia time constant
T_{DLC} DLC time constant
T_{PV} PV time constant
V_U, V_L Governor valve limiter
H_∞/H_2 H-infinity or H-2 controller
E_s Stored energy by ESS

1 Introduction

The invention of electricity in the 19th century changed the everyday habits of humankind. Nowadays it is difficult to imagine a world without constant electricity consumption—light, motors, modern transport, computers, communication—all these technologies have become viral in modern civilization. Nevertheless, every human activity has positive and negative outcomes. Consequently, fossil fuel-based energy has become one of the reasons for global warming, though trends have changed dramatically since international agreements on power have been signed by many countries. For the future of power engineering, an important step is to move towards the concept of a recyclable economy. Moreover, the International Energy Agency (IEA) predicts that by 2030, global electricity consumption will increase to 30,000 TWh, i.e. twice the energy consumed in 2010 [79].

These changes significantly increased demand for renewable energy sources, which should replace traditional power plants. However, the generation of stable and sustainable power from natural sources is a big challenge for today's infrastructure because AC power networks, widely used since their victory in the wars of the currents, were designed to be provided by power with stable generation and high inertia. In contrast, renewable power sources, including wind turbines and electronics-based solar panels, are unstable and have low inertia, the impact of which increases in parallel with degree of variable renewable energy (VRE) penetration. Therefore, VREs are applicable for grids designed to adapt generators with decreased inertia. An isolated microgrid is a great example of this power system, where a high-inertial thermal power plant (TPP) works alongside low-inertial wind generators and PV panels that have instant inertia drop. If inertia becomes too low, there is a high risk of a short-term grid shutdown and damage to many electrical devices. This can happen due to a drop in frequency and a high rate of change of frequency (RoCoF), hence the full scale of the impact so far requires deeper understanding [97, 61, 41, 46]. Therefore, low-inertia phenomena appeared due to the transition from a generator-dominated to an inverter-dominated power system (See Fig. 1). This poses a significant challenge to the industry of distributed power systems. So far, decreased inertia is an open problem that requires study balance models, stability theory and control methodologies, which still need to be developed and validated [61, 46, 36]. Moreover, microgrids introduce a complex and ramified network of interconnected generators with varying inertia and have many criteria for frequency stability support, which is critically important.
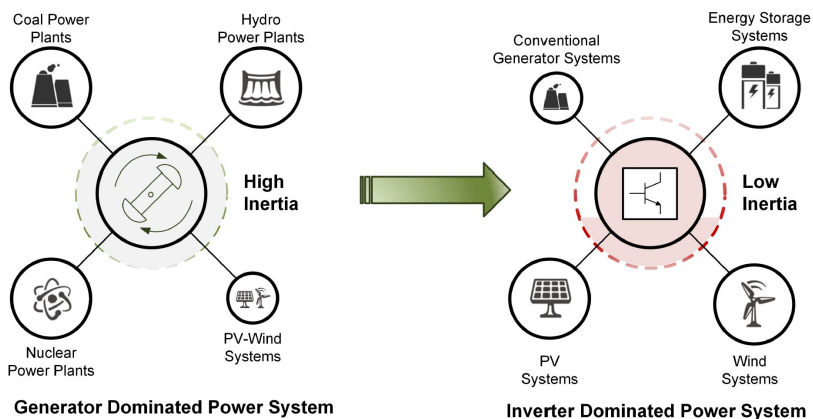

Figure 1: Evolution of inverter-dominated systems adopted from [93]

Nowadays engineers have proposed methodologies to facilitate the mitigation of problems. For instance, a common way is virtual inertia emulation, usually for inertia emulation, which applies the battery storage system combined with a power electronics device known as a virtual synchronous generator [98, 95, 61]. However, the inertia emulation process should be controlled using a robust method. In recent works, the droop control has been mentioned in [78, 24], where a comparative analysis between implementations for virtual synchronous machine were provided. Certain developments have been dedicated to frequency support by inertia response [13] and frequency measurement by phase lock-loop device (PLL) [93]. Some works propose inertia control provided by virtual inertia-based PI controller [28, 42] and use of the derivative control technique [9, 22], including the frequency support by method time-variable droop characteristic in [19].

The latest achievements in machine learning development (e.g. reinforcement learning, evolution algorithms and supervised learning) demonstrate high adaptivity to unstudied conditions in various tasks, such as medicine, economics, data mining, self-driving cars, cyber-security, gaming [20, 91]. Therefore, recent research focuses on searching for a solution based on computational intelligence. Recent developments in the field of artificial intelligence propose a large number of methods, but only a limited number of AI methods was applied for microgirds [110].

According to a recent study on the state of the art of the virtual inertia control methods provided in[86], several works tried to solve the problem of controllable virtual inertia emulation. All of the proposed methods have advantages and limitations that can have a serious impact on real implementation. For example, classical methods, such as the coefficient diagram method (CDM) and H_∞ in [2, 34] are robust but use the state-space representation of the process, which makes it difficult to recreate the dynamics of the real system and guarantee stability. Besides, H_∞ has problems with order reduction. The fuzzy logic controller (FLC) proposed in [38] is an advanced adaptive method with flexible architecture but requires good knowledge of the design of optimal fuzzy rules that perform manually alongside a computational complex. A hybrid method such as the optimal PI/PID controller tuned by particle swarm optimisation (PSO) in [53] is a realistic alternative, but learning takes place in the offline mode, which means that the optimal performance provided is limited and requires a long time for repeated tuning. Another hybrid method is the model predictive controller (MPC) proposed in [33, 90], which is a robust method that uses the finite-horizon method for adaption to uncertainties, but its performance heavily depends on the accuracy of the predicted input signal. Most of these methods are difficult to deploy for coordinated control or for another multi-scale system. Moreover, most of the methods lack any validation rules for fail-safe mode when applied to real processes in an unstudied environment.

This thesis is based on a collection of peer-revived scientific works published in reputed conferences and journals, where the study on virtual inertia problems and major contributions provided by the prepared doctoral thesis can be summarised as:

- **Design of ANN-based controller for virtual inertia control (VIC).** We design an ANN for VIC application and train by model-free reinforcement learning (RL) methods with DDPG and SDDPG policy rules.

- **Design of reward/punishment system for VIC applications.** We apply a data-driven approach for the study of unknown systems and train any controller by trial and error, where we design a specific set of reward and punishment rules to provide robust training of the RL agent in an unstudied environment and where the major criterion for rewarding is the calculated frequency stability.

- **Combination of advanced energy storage control scheme with MIMO neural network-based controller trained by deep reinforcement learning.** A redesigned ANN can be applied as a controller and/or an online tuning strategy for another control method. In an applied energy storage model we provide (dis)charge dynamics and optimal operation of a control scheme, where the designed MIMO controller organises control of incoming power flow and additional control of positive and negative feedback parts.

- **Design of multi-agent reinforcement learning control architecture for coordinated control of hybrid microgrid.** RL can be deployed as a multi-agent strategy and provide parallel training of multiple controllers. For example, a hybrid isolated microgrid model requires the design of a coordinated control strategy. We apply the multi-agent approach, which was trained by stochastic reinforcement learning (SRL).

- **Combination of neural network-based tuner with VFOPID controller.** The proposed hybrid model combines the robustness of VFOPID and the flexibility of an ANN, where training is supported by reinforcement learning. In addition, a combination of SRL and SDDPG policy is provided to optimise the duration of training.

1.1 Thesis structure

The thesis separated into seven chapters and covers one research topic, with two sub-topics of isolated microgrid and hybrid microgrid. A brief description of the content of each chapter is provided in the following list:

- **Chapter 1**, Introduction, provides an introduction to the topic virtual inertia control in isolated microgrid systems.

- **Chapter 2**, Background and challenges, addresses research questions and problems of the related topic.

- **Chapter 3**, Modelling of isolated microgrid, provides information on the features of the mathematical model isolated microgrid, including TPP, VIC, ESS, System inertia and damping, Domestic loads, VRE.

- **Chapter 4**, Analysis of methods for virtual inertia control in isolated microgrid, reflects the state of the art in recent algorithms applied for VIC, discusses their benefits and drawbacks.

- **Chapter 5**, Virtual inertia control in isolated microgrid mode, provides solutions on VIC proposed by the thesis, including a solution for the advanced energy storage model.

- **Chapter 6**, Coordinated frequency control in hybrid microgrid, provides a solution to the problem of the coordinated load-frequency control of hybrid microgrids.

- **Chapter 7**, Conclusion and Future Plans, addresses conclusion and final discussion of this thesis and future plans.

1.2 Research questions

In thesis we tried find answers for following research questions:

- How virtual inertia emulation help mitigate low inertia problem?

- What advantages reinforcement learning has in contrast to other optimization/training methods?

- What benefits combination of ANN and (FO)PID controllers has?

- How multi-agent reinforcement learning can be applied for coordinated control of hybrid microgrid?

2 Background and challenges

2.1 State of the art on low inertia systems

The low-inertia microgrid encompasses participants with different inertia power generation capacities and loads with very complex dynamics [28, 78, 24, 30]. Therefore, microgrids with high RES penetration are a big barrier for integration into massive distribution networks, creating various challenges such as: (1) active/reactive power imbalance and voltage droop in transmission lines; (2) production/consumption imbalance in distribution loads; and (3) frequency mismatch with other microgrids and the rest of the power grid [61, 59]. Hence, energy storage systems are considered as the prime actuator in frequency stability control, which in reality have physical limitations such as: (1) (dis)charge cycles; (2) restricted power reservation; (3) reserved power losses; and (4) individual speed of (dis)charge. Moreover, energy storage control performed by virtual inertia or virtual synchronous generator (VSG) uses power-inverting electronics, which come with physical delays and limitations in frequency measurement and power conversion [99, 108, 8, 107, 13, 93].

Renewable energy sources (RES) are frequently deployed in modern power grids to promote myriad environmental and economical aspects. However, increasing integration of RES significantly decreases the rotational inertia of the grid, which jeopardises grid stability and its overall dynamic behaviour [98, 41, 61, 42]. A central challenge is the regulation of the grid's frequency, considering the high penetration levels of renewable sources. One approach to the mitigation of this problem is the installation of fast-reacting storage systems with integrated virtual synchronous generators alongside low-inertia power sources. Such controllers have been studied extensively in recent years [9, 22, 101, 21, 69, 1, 36]. Each control method has its own benefits and limitations. For instance, classical control paradigms are simple in general but are designed for specific scenarios, whereas data-driven algorithms are flexible and enable online learning. On the other hand, these algorithms are numerically complex and require adequate data to operate efficiently. Furthermore, hybrid control strategies have low numeric complexity, but their convergence is hard to guarantee in most cases. Proposing suitable guidelines for choosing the best algorithm is currently an open question—a question which becomes more acute when the microgrid is isolated [93, 64, 109, 105, 53, 72, 36].

Isolated microgrids have received increasing attention as a means of integrating distributed generation into the electricity grid. Usually described as confined clusters of loads, storage devices and small generators, these autonomous networks connect as single entities to the public distribution grid through a point of common coupling (PCC). Fig. 2 illustrates a typical microgrid network. Microgrids comprise a variety of technologies: renewable sources such as photovoltaic and wind generators are operated alongside traditional high-inertia synchronous generators, batteries and fuel-cells. Thus, energy is generated near the loads, enabling the utilisation of small-scale generators that increase reliability and reduce losses over long power lines. The locality of the microgrid network enables the improved management of energy. Generators (and possibly loads) may be controlled by a local energy management system to optimise power flow within the network. The objectives of energy management depend on the mode of operation: islanded or grid-connected. In grid-connected mode, the typical objectives are to minimise the price of energy import at the PCC, to improve the power factor at the PCC, and to optimise the voltage/current profile within the microgrid. In islanded mode, which is addressed in this work, the main goal of power management is to stabilise the system and preserve high reliability and resiliency in terms of frequency, voltage and power.
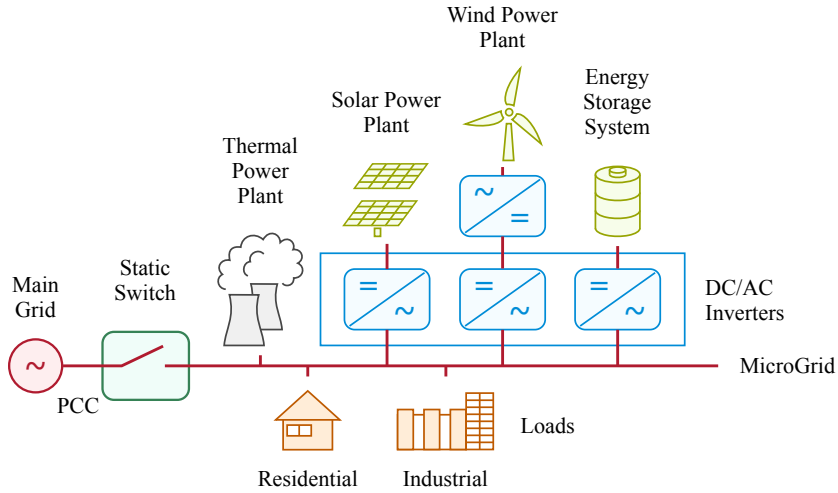
Figure 2: Schematic representation of a typical microgrid, including power conversion devices, power generators with different inertia

Few recent survey papers describe the various aspects in the context of virtual inertia within power grids with high penetration of RES. A comprehensive review of virtual inertia implementation techniques is provided in review work [93]. The reviewed works are classified and compared by means of virtual inertia topologies. Some selected topologies are simulated and it has been shown that a similar inertial response can be achieved by relating the parameters of these topologies through time and inertia constants. Moreover, a discussion on the challenges and research directions is presented and points out future research needs for the integration of virtual inertia systems. Review [82] presents various topologies for emulating the virtual inertia algorithm along with control strategies for general distributed generation. Furthermore, it provides a review of the optimal size and location of synthetic inertia in a power system. In [16], the authors presented a review focusing on inertia values for power systems. The inertia values were estimated based on different regions in the last 20 years. Furthermore, the contribution of PV power plants as virtual inertia is discussed and an analysis of the damping factor evolution provided.

Contrary to these comprehensive review papers, which focus on the implementation of virtual inertia topologies [93], virtual inertia and frequency control for distributed energy sources [82] and inertia estimation evolution in power systems [16], this study focuses on the systematic comparison of virtual inertia control methods designed to solve the frequency regulation problem in islanded microgrids. In particular, we strive to understand why certain control methods are more efficient in different circumstances and which control strategies will be gaining popularity in the upcoming years. To this end, we consider the different control techniques available in literature for the period from $2010-2023$ and then categorise them into three groups of *classic*, *advanced* and *hybrid* methods. Thereafter, we provide a detailed analysis of each control and the optimisation paradigms by means of various quality criteria. Finally, we perform a contextual analysis and highlight the current developments and trends for various combinations of virtual inertia control methods and technologies with a focus on microgrid applications.

2.2 Motivation

Complex and interconnected microgrids with decreased rotational inertia are a challenge for the development of new methods that can provide the effective injection of synthetic inertia via managing electronic-based devices, including DC/AC inverters, and stabilise energy in the multi-scale and decentralised energy generation network provided by grid-forming power sources, including thermal power plants, wind turbine parks and solar panel parks, as well as grid-following power sources, including ultra-capacitors, batteries, fuel cells, private solar panels and wind turbines. In simulated microgrid the frequency support provides by virtual inertia emulation, which in real life is a serious challenge for many classical and advanced control algorithms, since every microgrid has individual features and limitations that should be studied. In this thesis, the goal is to develop flexible algorithms based on artificial neural networks using the deep reinforcement learning approach. Many of the challenges are connected to the following tasks:

- Development of a reinforcement learning-based controller with (S)DDPG policy and design reward rules for application in a model of virtual inertia control.

- Study of microgrid dynamics to design the correct reward rules for training an RL control agent.

- Design of a robust artificial neural network-based controller frequency support/control in an isolated microgrid that can be a good alternative to the existing control methods proposed in the state of the art, including MPC, PID, FLC and H_∞.

- Application of the advanced energy storage scheme proposed in [47], which includes modelling of (dis)charge dynamics and limitations. Design of an ANN-based controller for applied system and development of simplified DDPG policy for control applications.

- Development of a hybrid of ANN and FOPID controllers for application in the virtual inertia controller scheme proposed in [36] in order to provide frequency support in an isolated microgrid.

- Development of multi-agent control architecture based on reinforcement learning to provide frequency support in a multi-area microgrid system.

- Development of multi-agent reinforcement learning for application in coordinated control of the hybrid microgrid model.

- Development of a reward system based on angular integral absolute error (IAE) to provide accurate punishing/rewarding of an RL agent.

3 Modelling of isolated microgrid

In this chapter, we illustrate the state of the art in modelling isolated microgrids applicable for testing control algorithms, which was proposed in [43] and in recent years saw an extension in an additional control loop provided by a virtual inertia controller. This thesis presents a simplified mathematical model of an isolated microgrid adopted from several recent publications [34, 38, 35, 2, 53, 84] and depicted in Fig. 3. The addressed scheme includes simplified residential/industrial loads, energy sources (thermal power plant, wind farm and solar power plant) and energy storage systems [76, 65, 36]. The thermal power plant is composed of a governor with a generator rate constraint (GRC) and a turbine with a frequency rate limiter, which restricts the valve opening/closing (V_U, V_L). The dynamic model of the microgrid utilises the hierarchical architecture with primary and secondary control loops. The primary control loop has a droop coefficient of $1/R$ and the secondary loop has an area control error (ACE) system with the second frequency controller with a gain of K_I and the first-order integrator. The frequency regulation is performed by the virtual inertia device with an additional controller. The balancing system performed as the first-order transfer function with microgrid damping coefficient D and system inertia H has a common value for all generators. The power generation by variable energy sources is modelled as a random signal with the first-order holder. The structure of hierarchical control includes frequency control by the primary control loop (PCL) and secondary control loops (SCL). The modelling parameters of the microgrid from different publications are summarised in Table 1.
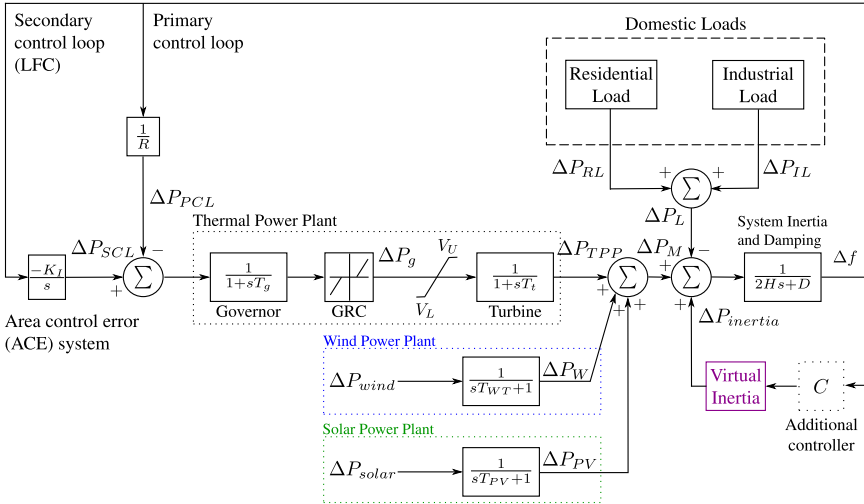
Figure 3: Mathematical model of isolated microgrid with hierarchical control loop and frequency support provided by virtual inertia controller, including renewable energy and domestic loads

3.1 Thermal power plant

In real life, a thermal power plant is represented as the major power supplier with a massive synchronous generator, which has high rotational inertia. In the applied mathematical model to control the frequency deviation Δf and preserve MG stability under various disturbances, two main frequency control schemes are employed: PCL, SCL. In this hierarchical architecture, PCL and SCL are responsible for balancing and restoring the system

Parameter	Physical meaning	Nominal value	Unit
ΔP_{TPP}	TPP power changes	–	p.u.
ΔP_E	General power deviations	–	p.u.
H	System inertia	0.083	p.u. s
D	Damping coefficient	0.015	p.u./Hz
ΔP_{VI}	Virtual inertia power changes	–	p.u.
K_{VI}	Virtual inertia constant	–	p.u. s
D_{VI}	Virtual damping coefficient	0.3	p.u./Hz
R_{VI}	Virtual inertia droop coefficient	2.7	Hz/p.u.
R	Droop coefficient	2.4	Hz/p.u.
T_{VI}	Virtual inertia time constant	10	s
T_t	Turbine time constant	0.4	s
ΔP_G	Governor power changes	–	p.u.
T_G	Governor time constant	0.1	s
ΔP_{SCL}	ACE action changes of SCL	–	p.u.
ΔP_{PCL}	Control changes of PCL	–	p.u.
K_I	Integral controller gain	0.075	s
ΔP_{RES}	RES power changes	–	p.u.
ΔP_{WTG}	WTG power changes	–	p.u.
ΔP_{wind}	Initial wind power variation	–	p.u.
T_{WTG}	WTG time constant	1.85	s
ΔP_{PV}	PV power changes	–	p.u.
ΔP_{solar}	Initial solar power variation	–	p.u.
T_{PV}	PV time constant	1.5	s
ΔP_L	Load power changes	–	p.u.
ΔP_{RL}	Residential loads variations	–	p.u.
ΔP_{IL}	Industrial loads variations	–	p.u.
V_U, V_L	Governor valve limiter	± 0.5	p.u.
$\Delta P_{g,\min}^{\max}$	Governor dead band limits	± 0.12	p.u. MW/min
$\Delta P_{VI,\min}^{\max}$	Virtual inertia valve limiter	± 0.25	p.u.

frequency stability, respectively. These two loops are applied to the thermal power plant (TPP) governor to generate power from the turbine system provided as

$$\Delta P_G = \frac{1}{1+sT_G}\left(\Delta P_{SCL} - \Delta P_{PCL}\right),\qquad(1)$$

in which $\Delta P_{PCL} = R^{-1}\Delta f$ and $\Delta P_{SCL} = s^{-1} - K_I\Delta f$ are the control and ACE action changes from PCL and SCL, respectively; R is the droop constant; and K_I is the integral controller

gain. Finally, its output is defined as:

$$\Delta P_{TPP} = \frac{1}{1+sT_t}\Delta P_G, \tag{2}$$

where ΔP_G is limited by valve opening/closing constants V_U and V_L.

3.2 Variable renewable energy sources

In traditional power plants, the synchronous machine has high inertia, the magnitude of which defines the total mass of the armature, so inertia is rather an advantage than a drawback. To understand the true scale of the problem, the fact that in traditional power plants the rotational speed of the generator depends on controllable steam flow, which is easily controlled by simple closed-loop methodologies, needs to be emphasised. In contrast, the kinetic energy of wind turbines is defined by wind velocity and air density, which is not controllable. Moreover, the rotational speed strongly depends on the total size and mass of the wind turbine. Unfortunately, future grids require an increase in wind turbine generation; an increased need for power requires larger and larger turbines. A similar situation occurs in relation to solar panels; however, photoelectric panels have zero inertia and energy generation depends on the daily intensity of solar radiation. In modelling, we consider variable renewable energy as a source with unstable power generation and decreased inertia, where solar panels ΔP_{PV} and wind turbines ΔP_{WTG} provided as the random signals with a transfer function that simulate natural transfer delay power flow. To make the simulated system more feasible, the renewable energy sources do not participate in frequency management and are considered MG uncertainties. Hence, the following simplified models of the renewable energy power changes are sufficiently accurate for our analysis:

$$\Delta P_{RES} = \Delta P_{WTG} + \Delta P_{PV} = \frac{1}{1+sT_{WTG}}\Delta P_{wind} + \frac{1}{1+sT_{PV}}\Delta P_{solar}. \tag{3}$$

where ΔP_{wind} and ΔP_{solar} are signals with random defined values, where T_{WTG} and T_{PV} are transfer delays of wind turbine and solar panels, respectively.

3.3 System inertia and damping

In the majority of power systems, the inertia model has a predefined constant of system inertia H (sometimes referred to as M), whose nominal value is 0.083 (i.e. 100% inertia) and can be changed to simulate the dynamics scenario with decreased inertia, e.g. with 80% and 40% of nominal magnitude; in the must-case load, damping coefficient D has a nominal value of 0.015. By defining the damping D and inertia constant H, the frequency deviation Δf can be represented as

$$\Delta f = \frac{1}{2Hs+D}\Delta P_E, \tag{4}$$

where ΔP_E is the general power deviation resulting from all power sources and loads and can be calculated as

$$\Delta P_E = \Delta P_{TPP} + \Delta P_{WTG} + \Delta P_{PV} + \Delta P_{VI} - \Delta P_L, \tag{5}$$

with $\Delta P_L = \Delta P_{RL} + \Delta P_{IL}$ as the load power changes, respectively. The system and damping function are based on a swing equation, which in the common form is represented as:

$$\frac{2H}{f}\frac{\mathrm{d}f}{\mathrm{d}t} = P_M - P_L \tag{6}$$

22

and simulates the simplified behavior of a real synchronous generator.

3.4 Energy Storage

Energy Storage is multidisciplinary domain in development technologies related to energy holding for a certain time. Since the invention of electricity, Energy Storage Systems (ESS) technologies have had implementations in various physical forms. The most perspective types of ESS are battery storage system, supercapacitor, flywheel, pumped hydrostorage and superconducting magnetic. During the last decade, energy storage technology became essential in power systems as never before [61, 76], as they allow the smoothness of any unstable energy source by energy accumulation for later dispatch. Energy storage can be directly incorporated into frequency response services and activated quickly, slowing down the RoCoF during a frequency event. Energy storage can be applied for any variable renewable energy generator, including Wind and Solar. Using this method, energy can be stored during overproduction and utilised during the underproduction of power in the grid. In fact, the majority of storage technologies are considered expensive developments that require additional enhancements. ESS has implementations in various physical realisations [14, 63, 76] and can be directly incorporated into frequency response services to support RoCoF during a frequency change event. For the last decade, ESS has become an essential component in the integration of variable renewable energy, since it may provide frequency smoothness and balance for further dispatch [44, 75, 106, 71, 5, 9, 96]. The simplified model of ESS can be represented as:

$$\Delta P_{VI} = \begin{cases} \Delta P_{VI\max}, & \Delta P_{VI} > \Delta P_{VI\max} \\ G(s)RoCoF, & \Delta P_{VI\min} < \Delta P_{VI} < \Delta P_{VI\max} \\ \Delta P_{VI\min}, & \Delta P_{VI} < \Delta P_{VI\min} \end{cases} \tag{7}$$

where $G(s)$ represented as first order function:

$$G(s) = \frac{1}{T_{VI}s + 1}, \tag{8}$$

where T_{VI} is time delay that simulates ESS speed and ΔP_{VI} the power injection limitations.

3.5 Domestic loads

Domestic loads simulate the dynamics of regular electricity consumers in every big settlement, including residential (i.e. residential buildings, civil institutions, private houses, schools) and industrial loads (i.e. factories, chemical plants, manufactures, mining institutions). This part of MG is reasonably considered as a variable set point in a simulated closed-loop isolated microgrid, where residential and industrial loads are summarised to one signal represented as:

$$\Delta P_L = \Delta P_{RL} + \Delta P_{IL}, \tag{9}$$

where ΔP_{RL} and ΔP_{IL} are randomly generated signals; to simulate natural delay, the first order holder is applied.

3.6 Virtual Inertia Control

A virtual synchronous generator (VSG) is a power converter-based device that produces the power alternative to the real synchronous machine [6, 23, 61]. A VSG is designed to compensate for the lack of inertia using a power injection mechanism. Such a generator

can be applied in systems with a high level of fluctuating renewable power to enhance the frequency stability. A virtual inertia controller is a simplified mathematical model of a VSG that neglects real measurements like voltage, current and (re)active power control [97, 61, 36]. The default operational limitations of a virtual inertia controller (VIC) cannot provide reliable frequency support. Therefore, the additional robust controller has to be used to deal with nonlinearities in low-inertia environments. Traditionally, a virtual inertia control setup (see Fig. 4) consists of a derivative component, virtual inertia variable gain K_{VI}, an energy storage system and a power limiter ($\Delta P_{VI,\max}$, $\Delta P_{VI,\min}$). The main concept used in virtual inertia control is the so-called Rate of Change of Frequency (RoCoF), which can be calculated as:

$$RoCoF = \frac{\mathrm{d}(\Delta f)}{\mathrm{d}t}. \tag{10}$$

The RoCoF defines the time derivative of the frequency signal, which is used to calculate the inertia response of the system as

$$\Delta P_{VI} = \frac{K_{VI}}{1 + sT_{VI}} \frac{\mathrm{d}(\Delta f)}{\mathrm{d}t}, \tag{11}$$

where virtual inertia constant K_{VI} is usually defined as

$$K_{VI} = \frac{2HP_{Inv}}{f_0}, \tag{12}$$

where f_0 is the nominal frequency, P_{Inv} is the power of the inverter and H is the calculated inertia [53, 2, 37].

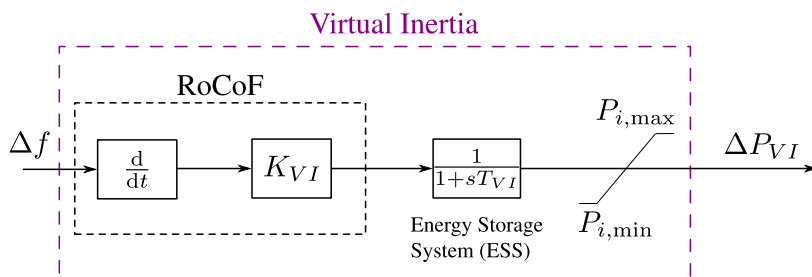Figure 4: Typical structure of virtual inertia controller with constant K_{VI} described as virtual inertia gain

3.7 Enhanced virtual inertia controller

An enhanced virtual inertia controller can be provided by:

$$\Delta P_{VI} = \frac{sK_{VI} + D_{VI}}{1 + sT_{VI}} \left(\frac{\Delta f(s)}{R_{VI}} \right). \tag{13}$$

The enhanced VIC proposed by [36] decreases the influence of decreased inertia (see 5), which provides an illustrative comparison between versions of virtual inertia controller, illustrated in Fig. 4. (a)—full scheme with proposed by [36]; (b)—scheme without virtual droop block $1/R_{VI}$ proposed by [37]; and (c)—scheme without virtual droop $1/R_{VI}$ and virtual inertia damping D_{VI} proposed by [34, 53, 2, 84]. In the last version, we can see the significant effect of decreased inertia. To be more specific, the minimum level of inertia, or its critical floor, is usually used to determine the minimum inertia required for

the system to be stable and operate in a safe mode. This level can be determined based on the maximum $RoCoF$ and the frequency nadir. These two constraints can be determined by analysing the transient response of the system. As can be seen in Fig. 5, all three results show different $RoCoF$ and frequency overshoot and $nadir$ for different levels of inertia. When the system inertia and damping are reduced due to renewable energy sources and distributed generations penetrations, the constraint nadir/overshoot of the microgrid significantly increases, resulting in a longer stabilising time. Additionally, the RoCoF of the system increases, meaning that a system operator has less time to respond to disturbances. In this case, following the disturbance, the generated power fluctuates more, leading to the stress in this unit.
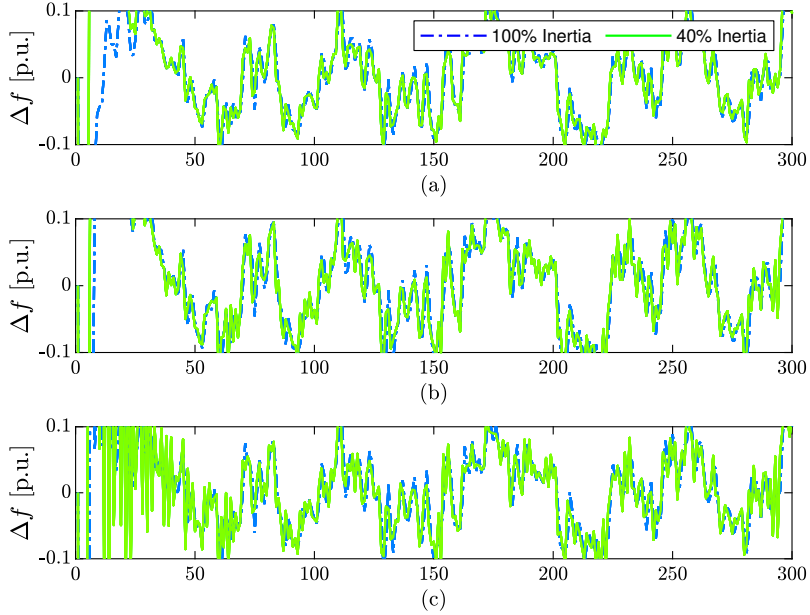

Figure 5: Illustration of decreased inertia effects with variations of virtual inertia controller: (a)—full scheme with constant K_{VI}; (b)—without $1/R_{VI}$ and (c)—without $1/R_{VI}$ and D_{VI}.

4 Analysis of methods for virtual inertia control in isolated microgrid

Several recent works [93, 52, 34, 53, 2, 38, 37, 35] address the problem of optimal frequency support with high penetration of variable renewable energy sources. For example, in [35], a robust H_∞ controller was designed to provide stability support based on the Rate of Change of Frequency (RoCoF). The proposed solution has shown advantages over conventional virtual inertia control and optimally tuned PI controllers in scenarios where the wind farm is connected, solar panels are disconnected and the system inertia is 10% and 100%, respectively. In [38], the problem was further studied by implementing a virtual inertia control scheme combined with a fuzzy-logic based approach. The proposed algorithm performed robustly under different scenarios with additional uncertainties, including 80%, 40% and 30% total system inertia and a mismatch in primary/secondary control loops. In [33], the model predictive control scheme was proposed and compared to the fuzzy-logic controller for the case of additional loads connection. Unlike the previous works, the studied microgrid has conceptual differences, such as a closed-loop TTP system, RES power generation from two complex wind farms and minor differences in transfer function describing turbine and system inertia. Similar ideas were presented in [92] but without modelling renewable energy disturbances. Work [53] presents a PI controller optimised by particle swarm optimisation and combined with the digital frequency protection system in scenarios of (dis)connecting load and renewable energy sources.
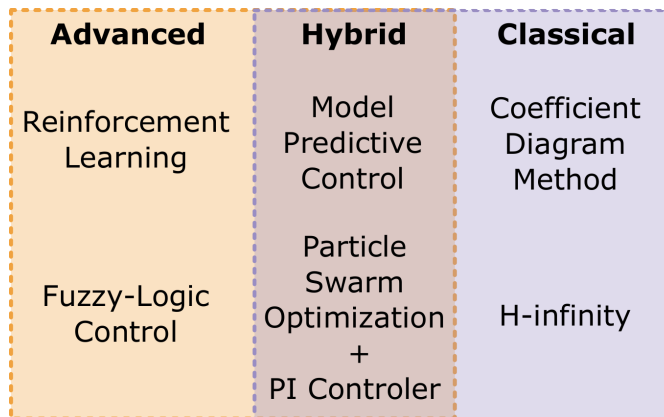
Advanced	Hybrid	Classical
Reinforcement Learning	Model Predictive Control	Coefficient Diagram Method
Fuzzy-Logic Control	Particle Swarm Optimization + PI Controler	H-infinity

Figure 6: Classification of algorithms for virtual inertia control

4.1 H-infinity

H_∞ applies the synthesis of an optimal controller by taking into consideration microgrid disturbances and uncertainties via state-space representation, which can provide high robustness and simple hardware realisation. However, the main difficulty is the necessity to design an accurate state-space description for tuning the controller [34, 35]. The frequency control based on H_∞ was used in [83, 39, 15, 34, 35]. In [104], mixed H_2/H_∞ was applied for multi-area microgrid and compared to fuzzy logic controllers. The solution presented in [35] applies linear fractional transformation in optimal H_∞ regulator design as the basis for modelling microgrid uncertainties z, such as system inertia H, damping properties D and PLL delays (ω_n and ζ). H_∞ optimisation preforms in offline mode and is more vulnerable to low-inertia nonlinearities than data-driven algorithms. At the same

time, synthesis of the robust model by H_∞ provides reliable frequency support. For example, work [34] shows the implementation of this method, successfully tested with 95%, 45% and 15% of the nominal system inertia and using two types of disturbance: $1)$ 10% of step change in load power demand; and $2)$ a mismatch in microgrid generation by increased time constant of governor and time constant of turbine. Work [35] shows the H_∞ controller tested with 100% and 10% of system inertia in scenario with 80% renewable energy penetration. However, the common limitation of the H_∞ method is notable peaks during (dis)connection of power plants. H_∞ requires detailed understanding of classical control theory and optimisation, which does not require powerful hardware for operation. Nevertheless, the synthesised control model is a high-order transfer function and often requires order reduction [34, 35]. The biggest difficulty of H_∞ optimisation is the procedure for the development of an accurate state-space representation and the manual estimation of disturbances. A controller based on the coefficient diagram method (CDM) relies on an algebraic optimisation approach through polynomial state-space representation and the Routh-Hurwitz stability criterion [56, 57], where the theoretical basis is constituted to satisfy the Lipatov and Sokolov stability criterion [49, 55, 4, 12].

4.2 Coefficient Diagram Method

Similarly to H_∞, the optimisation procedure is designed for offline mode. Implementation of a controller based on CDM in [2, 3] provides frequency stability in a range less than ± 0.1 Hz in a scenario with 100% inertia and two types of disturbance: (1) 10% step load perturbation; and (2) random load demand. In contrast to H_∞, it can mitigate peaks after (dis)connection of renewable energy sources. However, the solution uses a two degree of freedom system structure expressed as $N(s)/D(s)$, which is designed to track a limited number of disturbances. The main drawback of CDM controller synthesis is similar to that of H_∞ and relies on a good understanding of classical control theory optimisation. At the same time, it can be implemented using relatively simple hardware. In contrast to H_∞, CDM optimisation performs without necessity in order reduction and utilises the coefficient method instead of the Bode diagram [2, 3]. However, the validation of synthesised control by the Routh-Hurwitz or Lipatov-Sokolov stability criterion depends on the order of the synthesised control system [4, 12, 2, 3].

4.3 Fuzzy Logic Controller

The fuzzy logic controller design provides effective manual optimisation compared to other advanced algorithms. Several examples of frequency regulation can be found in [62, 7, 77, 60, 29, 38, 10]. Since a fuzzy logic-based controller can be manually tuned, the data-driven approach is optional. Correct configuration of the controller can make a robust system. Work [38] applies a standard fuzzy logic controller for virtual inertia control, which was mostly capable of holding Δf inside a ± 0.1 Hz band with 80%, 60% and 30% system inertia in scenarios with 20%, 80% RES penetration and a mismatch in primary/secondary control loops. Controller design requires a good understanding of the design principles of fuzzy rules. In addition, this method requires powerful hardware for implementation. However, it utilises fuzzy logic without an optimiser, which can be considered a drawback, since it requires optimal fuzzy rules by manual design [54, 60, 104].

4.4 Evolutionary Optimisation

Particle swarm optimisation is a popular evolutionary algorithm inspired by collective species behavior, such as a flock of birds [31], the stochastic optimisation of which should provide the best performance via searching a global minima. Particle swarm strategy is a stochastic data-driven optimiser and enables online learning [18, 32, 53, 51]. In work [53], PSO is used for the optimal tuning of PI via searching the global minima of a microgrid and provides robust control with 100%, 80% and 30% system inertia. Performance of the optimal PI demonstrated in [53] shows relatively stable frequency support with 100%, 50% and 30% system inertia and 57% RES penetration. In contrast to other solutions, in [53], a dynamic model of a microgrid with digital protection is applied and provides additional frequency stability. PI/PID is a widely used controller in the power industry since its construction is simple [74, 17, 27]. However, the PSO algorithm belongs to the self-learning optimiser, whose implementation is more complex. In order to receive the optimally tuned PI controller, the optimiser has to consider the state-space dynamic modelling of microgrid uncertainties, which requires a relatively long time for finding the optimal settings, which in turn is not safe for industrial applications.

4.5 Model Predictive Control

A model predictive controller (MPC) requires the development of a robust prediction model based on a detailed representation of process dynamics via collected data or state space [11, 58, 33]. As a hybrid algorithm, the MPC can be implemented with a data-driven [26] or finite time-horizon [40, 90] optimisation approach. Work [33] shows the application of the finite impulse response optimisation for model prediction based on the virtual inertia emulation with microgrid state-space representation. In aspects of optimisation, the MPC can provide real-time learning using a data-driven and finite-horizon approach. According to [33], the control performance of the MPC is higher than that of the fuzzy logic-based controller and may provide better Δf stability during (1) (dis)connection RES power; (2) sudden load change; and (3) mismatch in the main thermal generation scenarios with 100%, 50% and 25% system inertia and 34% RES penetration. The implementation of a model prediction-based controller depends on the type of prediction model. The controller requires calculation of each time sample and heavily depends on the designed model used in predictions of microgrid disturbances [33, 92]. Specifically, in [33] the finite impulse response was used, which considers each sampling instantly in the prediction of microgrid disturbances.

4.6 Reinforcement learning-based controller

Reinforcement learning (RL) is an agent-based and model-free machine learning algorithm [91]. The main feature of RL optimisation is an approach based on trial and error, which allows validation of the artificial neural network (ANN)-based controller directly with the control object and prediction of any negative consequences [48, 91, 84]. The benefit of this method is mandatory data-driven optimisation, which is naturally designed for online learning. In [84], RL was compared to H_∞ and resulted in the slightly better performance of frequency stability in scenarios with 100%, 80% and 40% inertia and connection of wind, solar and thermal plants during the launch of industrial and residential loads, and 20% RES penetration. In [86], frequency stability with 50% RES penetration was provided. Since the algorithm uses a deep neural network, it requires strong computational hardware and is relatively complex to implement. The method requires the selection of optimal action $a(t)^*$ at each step $s(t)$ and takes a long time. Moreover, for RL,

it is necessary to design a proper reward system and choose the right training strategy, which can differ [25, 80, 103, 84]. For example, in works [25, 80], the RL optimisation for frequency support was performed by approximated dynamic programming. In contrast, work [84] uses the deep deterministic policy gradient to train an RL-based controller for virtual inertia emulation.

4.6.1 Neural Actor-Critic Architecture

Neural actor-critic implies the combination of two Artificial Neural Networks: the actor network $\mu(s \mid \theta_\mu)$; and the critic network $Q(s, a \mid \theta_Q)$. In the proposed strategy, the critic network supervises the algorithm, which tracks errors from interactions between actor a_t and environment S_t according to the defined policy. The network corrects them in order to find the optimal estimation of actor action a'_t, which predicts the maximum possible reward r (see [91]). The key advantage of reinforcement learning algorithms is the study based on interaction with the environment [91]. This means that when an agent makes the action, it expects to get the reward $+r$ or the punishment $-r$. The mechanism of control can be briefly summarised as follows: The measured frequency difference Δf produces the control error, which goes as an observation to the RL agent. At the same time, the calculated error goes to the block "calculate reward" to reward or punish the neural RL actor.

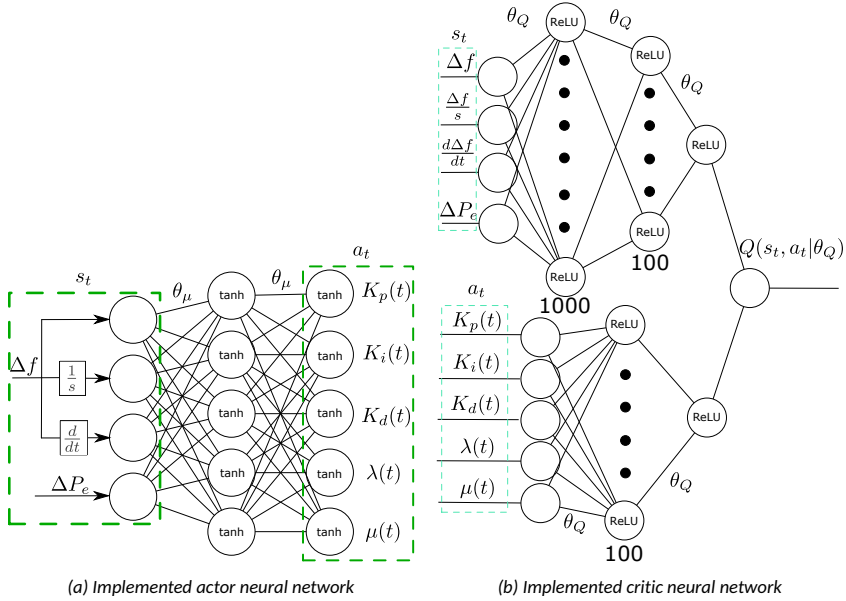(a) Implemented actor neural network   (b) Implemented critic neural network

Figure 7: Example of implemented neural networks of actor (i.e. Tuner) and critic with illustrated connections between them defined as action a_t and state s_t (green dashed line), applied in publication [67], where Δf, $\Delta f/s$, $d\Delta f/dt$ is state s_t and $K_p(t)$, $K_i(t)$, $K_d(t)$, $\lambda(t)$, $\mu(t)$ is action a_t

4.6.2 Deep deterministic gradient descent

Deep Deterministic Policy Gradient is a reinforcement learning algorithm designed for tasks with a low-dimensional continuous action space [48]. The DDPG optimisation introduces the fusion of deep Q-learning (DQL) and deterministic policy gradient (DPG), which inherits neural actor-critic from DQL strategy. The principle of optimisation is based on the search of a minimal difference between the target action-value function of the policy y_i and the critic network $Q(s_i, a_i \mid \theta_Q)$ at each actor network $\mu(s \mid \theta_\mu)$ decision a_t (K_{VI}) per step i in the state s_t (i.e. Δf_t) in order to minimise the loss function L and receive the maximum possible reward r_t per training episode. Next, we summarise the overall procedure in the form of a pseudo-code as shown in Algorithm 1, adapted from [48].

Algorithm 1 DDPG Algorithm

1: Initialise critic $Q(s, a \mid \theta_Q)$ and actor $\mu(s \mid \theta_\mu)$ networks with random weights θ_Q and θ_μ.
2: Initialise target network Q' and μ' with $\theta_{Q'} \leftarrow \theta_Q$, $\theta_{\mu'} \leftarrow \theta_\mu$.
3: Initialise replay buffer R.
4: **for** $episode = 1$ to M **do**
5:      Receive initial process observation as state s_1.
6:      **for** $t = 1$ to T **do**
7:          Select action $a_t = \mu(s_t \mid \theta_\mu)$ according to current policy and disturbances exploration.
8:          Execute action a_t. Observe reward r_t, state s_{t+1}.
9:          Store transition (s_t, a_t, r_t, s_{t+1}) in R.
10:         Sample random minibatch of N transitions (s_i, a_i, r_i, s_{i+1}) from R.
11:         Set $y_i = r_i + \gamma Q'(s_{i+1}, \mu'(s_{i+1} \mid \theta_{\mu'}) \mid \theta_{Q'})$.
12:         Update critic by minimising loss: $L = \frac{1}{N} \sum_i (y_i - Q(s_i, a_i \mid \theta_Q))^2$.
13:         Update actor policy using sampled gradient:
14:         $\nabla_{\theta_\mu} J \approx \frac{1}{N} \sum_i \nabla_a Q\left(s, a \mid \theta^Q\right)\big|_{s=s_i, \alpha=\mu(s_i)} \times \nabla_{\theta_\mu} \mu(s \mid \theta^\mu)\big|_{s_i}$.
15:         Update the target network: $\theta^{Q'} \leftarrow \tau\theta^Q + (1 - \tau)\theta^{Q'}$ and $\theta^{\mu'} \leftarrow \tau\theta^\mu + (1 - \tau)\theta^{\mu'}$.
16:      **end for**
17: **end for**

4.6.3 Simplified Deep Deterministic Policy Gradient

In contrast to the original DDPG, the modified version has a simplified loss function for actor L_μ and training is performed without a replay buffer. The principle of optimisation is based on the search for a minimal difference between target action-value function y_i and critic network reward prediction $Q(s_i, a_i \mid \theta_Q)$ for each actor network $\mu(s \mid \theta_\mu)$ decision a_t (i.e. K_{VI}) per step i in microgrid state s_t (i.e. Δf_t). The aim of the simplified DDPG algorithm is to receive the maximum possible average reward r_t per training episode via minimisation of critic L_Q and actor L_μ loss functions. The pseudo-code shown in Algorithm 2 illustrates the simplified DDPG algorithm.

Algorithm 2 Simplified DDPG algorithm

1: Initialise critic $Q(s, a \mid \theta_Q)$ and actor $\mu(s \mid \theta_\mu)$ networks with random weights θ_Q and θ_μ.
2: Initialise learning rate for critic $\alpha_Q = 0.1$ and actor $\alpha_\mu = 0.1$ networks.
3: Initialise smooth $\tau = 0.5$ and discount factor $\gamma = 0.25$.
4: **for** $episode = 1$ to M **do**
5:    **for** $t = 1$ to T **do**
6:        Receive initial process observation as state s_t.
7:        Select action $a_t = \mu(s_t \mid \theta_\mu)$ according to current reward prediction $Q(s_i, a_i \mid \theta_Q)$.
8:        Execute action a_t, observe reward r_t and future state s_{t+1}.
9:        Select the best critic reward prediction weights $Q'(s_{i+1}, \mu'(s_{i+1} \mid \theta_{\mu'})$ according to explored max average reward $\max(\frac{1}{N} \sum_i r_i(s_i, a_i))$
10:        Set action-value function of DDPG policy $y_i = r_i + \gamma Q'(s_{i+1}, \mu'(s_{i+1} \mid \theta_{\mu'}) \mid \theta_{Q'})$.
11:        Update critic by minimising the loss: $L_Q = \frac{1}{N} \sum_i (y_i - Q(s_i, a_i \mid \theta_Q))^2$.
12:        Update actor policy using sampled gradient of critic: $L_\mu = \frac{1}{N} \sum_i (Q(s, a \mid \theta_Q) \mid_{s=s_i, \alpha=\mu(s_i)})^2$
13:        Update target networks:
14: $\theta^{Q'} \leftarrow \tau \theta^Q + (1 - \tau) \theta^{Q'} \alpha_Q L_Q,$
15: $\theta^{\mu'} \leftarrow \tau \theta^\mu + (1 - \tau) \theta^{\mu'} \alpha_\mu L_\mu.$
16:    **end for**
17: **end for**

4.6.4 SDDPG training algorithm with semi-stochastic approach

We apply a basic stochastic RL (SRL) combined with the SDDPG technique to train an agent implemented as trainer systems for ANN-based control of a microgrid. In the first stage of the proposed algorithm, the provided SRL algorithm m times generates weights with a random value and verifies every combination of the obtained weights through a "reward" feedback mechanism.

According to the proposed SRL optimisation approach, the RL agent will be defined as $\mu(s \mid \theta_\mu)$ with an initialised random set of weights $\sum_{i=1}^{M} \theta_{\mu_i}$, where θ is the array of non-negative random values that updates the defined number of episodes M. Then, the reward function $r_t(a_t, s_t)$, where $a_t = \Delta K_{VI}(t)$ and $s_t = \Delta f$, will be observed at each action and state. Finally, the weights θ_μ of the agent with the best average reward will be selected using

$$\theta_{\mu_l} = \begin{cases} \theta_{\mu_{M+1}} = \theta_{\mu_M}, & \text{if } r_l > r_{\max}, \\ \theta_{\mu_M} = \theta_{\mu_M}, & \text{if } r_l < r_{\max}, \end{cases} \tag{14}$$

where $r_{\max} = \operatorname{argmax}(r_t(s_t, a_t))$. The pseudo-code of this optimisation approach, applied to the first stage, is summarised in Algorithm 3.

Algorithm 3 SRL Optimisation Approach.

1: Initialise neural network of agent as $\mu(s \mid \theta_\mu)$
2: **for** $j = 1$ to M **do**
3:     Initialise random weights of each agent $\theta_{\mu_{l,M}}$ M times
4:     Execute action a_t, observe the reward r_t at each step t in common state s_t
5:     **if** Maximum reward $r > r_{\max}$ **then**
6:         Save weights of each agent θ_μ
7:     **end if**
8: **end for**

In the second stage of learning, we use the SDDPG method to improve the performance of the proposed algorithm. To this end, the gradient descent, which acts in a similar way to Q-learning architecture Actor-Critic with multiple agents, is applied. The proposed Actor-Critic with SDDPG technique provides gradient descent for each ANN-based controller in order to train them simultaneously. We design the reward system for each agent and take into account the common calculated reward. This stage is more computationally complex and provides the additional changes in the received weights from stage one by SDDPG (which inherits major features of the original DDPG method, proposed in [48]). In this stage, and by using the SDDPG technique, we try to receive the maximum possible reward formulated as $r_{\max} = \operatorname{argmax}(r_i(s_i, a_i))$ via reward prediction, which is provided by the output of the critic ANN given as $Q(s_t, a_t)$ and obtained by

$$y_t = r_t + \gamma Q'(s_{t+1}, \mu'(s_{t+1}|\theta_{\mu'}) \mid \theta_{Q'}), \tag{15}$$

and comparing the predicted reward with the actual one. The accuracy of reward prediction depends on the quality of monitoring in the actor's ANN state s_t and action a_t, which is defined by the critic weights θ_Q. In the following, the algorithm changes the weights of the actor's ANN (i.e. θ_μ) via gradient descent and tests an ANN-based controller after every change in the weights.

In general, control applications are not easy to solve with any machine learning method. This is especially true when training an ANN-based controller, since many gradient descent methods, such as supervised learning, take a long time to find the optimal weights

for such controllers. As the first stage is more effective for using stochastic methods (e.g. dynamic programming, brute force method), we have 20 episodes of training to find optimal weights, while it takes at least 200 episodes in the second stage associated with the SDDPG technique. Nevertheless, the gradient descent and applied policy increase the performance of the proposed controller by up to 5–15%. In both stages, validation of the controller is performed by the reward system. The following equation demonstrates how the DDPG policy affects the quality of the prediction in the reward system defined as action-state function $Q(s_t, a_t)$:

$$L_Q = \frac{1}{N} \sum_t (y_t - Q(s_t, a_t \mid \theta_{Q_i}))^2. \tag{16}$$

This equation is the full form of the loss function, including the critic's prediction of the reward system. The output of the critic provides the change in the actor weights θ_μ by

$$L_{\mu_i} = \frac{1}{N} \sum_t (Q_i (s, a \mid \theta_{Q_i}) \mid_{s=s_t, \alpha=\mu(s_t)})^2. \tag{17}$$

In order to increase the quality of the prediction algorithm, we store the best prediction as $Q'(s_{t+1}, \mu'(s_{t+1} \mid \theta_{\mu'}) \mid \theta_{Q'})$.

All described steps of method summarised in Algorithm 4.

Algorithm 4 SDDPG algorithm with semistochastic method.

1: Initialise critic $Q(s, a \mid \theta_Q)$ actor $\mu(s \mid \theta_\mu)$ networks
2: **for** $t = 1$ to N **do**
3:     Initialise random weights of actor θ_μ.
4:     Execute action a_t, observe reward r_t at each step
5:     **if** $r_{avg} > r_{avg\_max} = \max(\frac{1}{N}\sum_i r_i(s_i, a_i))$ **then**
6:         Save weights θ_Q and θ_μ with best r_{avg}.
7:     **end if**
8: **end for**
9: Initialise random weights of critic θ_Q.
10: Define the learning rate of critic α_Q and actor α_μ networks.
11: Define the smooth factor constant $\tau \in [0, 1]$.
12: **for** $episode = 1$ to M **do**
13:     **for** $t = 1$ to T **do**
14:         Receive initial process observation as state s_t.
15:         Select action $a_t = \mu(s_t \mid \theta_\mu)$ according to current reward prediction $Q(s_i, a_i \mid \theta_Q)$.
16:         Execute action a_t, observe reward r_t and future state s_{t+1}.
17:         Select the best critic reward prediction weights $Q'(s_{i+1}, \mu'(s_{i+1}|\theta_{\mu'})$ according to the explored r_{avg_max}.
18:         Set the state-action-value function of DDPG policy $y_i = r_i + \gamma Q'(s_{i+1}, \mu'(s_{i+1}|\theta_{\mu'}) \mid \theta_{Q'})$.
19:         Update the critic by minimising the loss: $L_Q = \frac{1}{N}\sum_i (y_i - Q(s_i, a_i \mid \theta_Q))^2$.
20:         Update actor policy using the sampled gradient of critic: $L_\mu = \frac{1}{N}\sum_i (Q(s, a \mid \theta_Q)|_{s=s_i, \alpha=\mu(s_i)})^2$
21:         Update target networks:
22: $\theta^{Q'} \leftarrow \tau\theta^Q + (1-\tau)\theta^{Q'}\alpha_Q L_Q,$
23: $\theta^{\mu'} \leftarrow \tau\theta^\mu + (1-\tau)\theta^{\mu'}\alpha_\mu L_\mu.$
24:     **end for**
25: **end for**

4.7 Reward system for training of RL agent

4.7.1 Error band-based reward system

To ensure the learning of the agent, we use the reward/punishment system proposed in [84]. The reward system is organised in such a way that the measured frequency deviation is converted to reward/punishment $r_t\pm$. To provide the instructions according to the stability criteria of error signal Δf in the ± 0.1 Hz band [84, 87], the regulation system is organised as follows: if $|\Delta f| < 0.05$, the agent receives a reward $r+$; otherwise, the system punishes $r-$ after each performed action. To provide a reasonable reward for the agent, we separated the rewarding approach for each task; the frequency support reward task is limited to the range $u \in \{1.8, \ldots, 2\}$ and the punishment is unlimited and multiplied by 2. The following equation presents the mathematical expression for the designed rules:

$$r_t = \begin{cases} \frac{1}{0.5+|\Delta f|}, & \text{if } |\Delta f| < 0.05, \\ -2|\Delta f|, & \text{if } |\Delta f| > 0.05. \end{cases} \tag{18}$$

4.7.2 Error angle-based reward system

The proposed reward system was applied in [88], the main criterion here for rewarding is the magnitude measured state $s_t = \Delta f$, which is considered as the measured control error calculated as integral absolute error (IAE). The agent receives a positive reward $+r$ if φ is less than 45 deg; otherwise, it receives a negative reward or punishment $-r$. In contrast to [84], the new reward system is more adaptive for control applications associated with high dynamical changes, such as interconnected microgrids, since it is based on a performance metric (IAE), which is a good criterion for process control. The major problem with multiple agents is synchronised training because the actions of the first agent affect the results of the other two agents and *vice versa*. Therefore, we admit that the reward of agents is easier to individually be calculated by a common summarised reward and save the weights with the best reward for each agent. The rewarding rules described can be summarised as follows:

$$r_t = \begin{cases} 2\frac{\text{IAE}}{\sqrt{\text{IAE}^2+t^2}}, & \text{if } \sin\varphi < 0.5 \\ -10\frac{t}{\sqrt{\text{IAE}^2+t^2}}, & \text{if } \sin\varphi > 0.5, \end{cases} \tag{19}$$

where IAE denotes the integral absolute error calculated by

$$\text{IAE} = \int_0^t |\Delta f| \mathrm{d}t. \tag{20}$$

4.8 Conclusion

This chapter reviews recent methods related to virtual inertia control methods designed to solve the frequency regulation problem in islanded microgrids, with an attempt to better understand the unique characteristics, common uses and mathematical foundations of the most popular control methods. This analysis reveals interesting trends in the current research and may help to understand why certain control methods are more efficient in different circumstances and which control strategies will be gaining popularity in the coming years. For instance, the data show that evolutionary algorithm methods are widely used for tuned PI/PID controllers, likely because this enables analysis of stochastic scenarios with nonlinear constraints. However, evolutionary algorithms may converge to the local minimum and therefore are not suitable for every application. In such cases, classical control methods seem to be the natural choice, since they provide simple and effective solutions to virtual inertia problems whenever grid dynamics are well-defined. Moreover, in the case of uncertainty in the grid dynamic and nonlinear constraints, fuzzy logic-based controllers are used extensively, although these fuzzy controllers are limited to specific and manually defined rules and in the case of high number of rules the required resources increase significantly. The controller based on the coefficient diagram method principle seems to be the least popular method, perhaps due to its limitation of tracking a limited number of disturbances. Artificial neural networks in combination with reinforcement learning are also becoming an important trend due to increasing amounts of available data—specifically, these model-free approaches can be used for solving complex problems when no fully satisfactory algorithm is available.

5 Virtual inertia control in isolated microgrid model

In this chapter, we discuss the control methods applied for virtual inertia emulation in isolated microgrids proposed in the thesis and illustrated in Fig. 3 with different versions of VIC, whose proposed implementations were studied in publications [84],[86] and [89].

5.1 Reinforcement learning-based controller

In [84], neural actor-critic implies the combination of two ANNs: the actor network $\mu(s \mid \theta_\mu)$, and the critic network $Q(s, a \mid \theta_Q)$. In the proposed strategy, the critic network is supervising the algorithm, which tracks errors from the interaction of the actor a_t with the environment s_t according to the defined policy. The network corrects them in order to find an optimal estimation of actor action a_t', which predicts the maximum possible reward r (see [91]). The key advantage of reinforcement learning algorithms is the study based on interaction with the environment [91, 48]. This means that when an agent makes the action, it expects to get the reward $+r$ or the punishment $-r$. The mechanism of control can be briefly summarised as follows. The measured frequency difference Δf produces the control error, which goes as an observation to the RL agent. At the same time, the calculated error goes to the block "calculate reward" to reward or punish the neural RL actor. The general structure of the proposed controller is presented in Fig. 8.
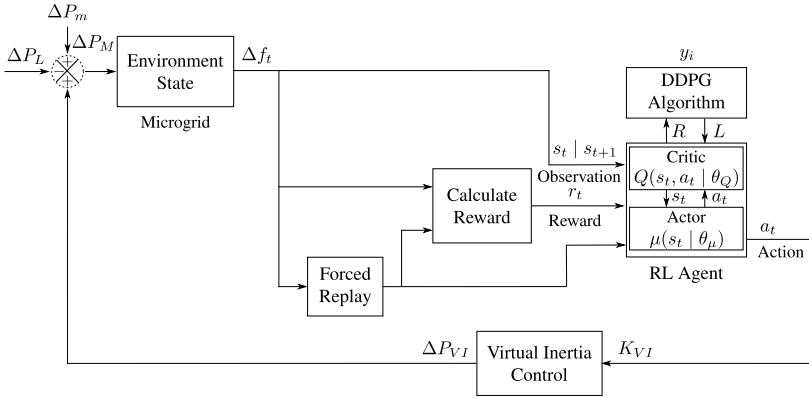

Figure 8: Structure of implemented reinforcement learning-based controller developed in [84], including the forced replay block, applied to decrease number of episodes with negative reward

5.1.1 Numerical results

Figure 10 illustrates the case without additional virtual inertia control; all control algorithms provide almost identical results. This is further emphasised by the statistical measures shown in Table 2. The proposed algorithm and H_∞-based approach both appear to be the most accurate. However, their advantage is rather symbolic with respect to PI and default controller cases. Therefore, in systems with high inertia, utilisation of additional virtual inertia controllers is rather optional. To conclude, the proposed reinforcement learning-based additional virtual inertia controller appears to be more accurate in most cases. However, the difference in most cases is very small. At the same time, it can be observed that on average the proposed controller performs better than other tested controllers as defined by mean average error. This is further confirmed by another related integral absolute error measure.
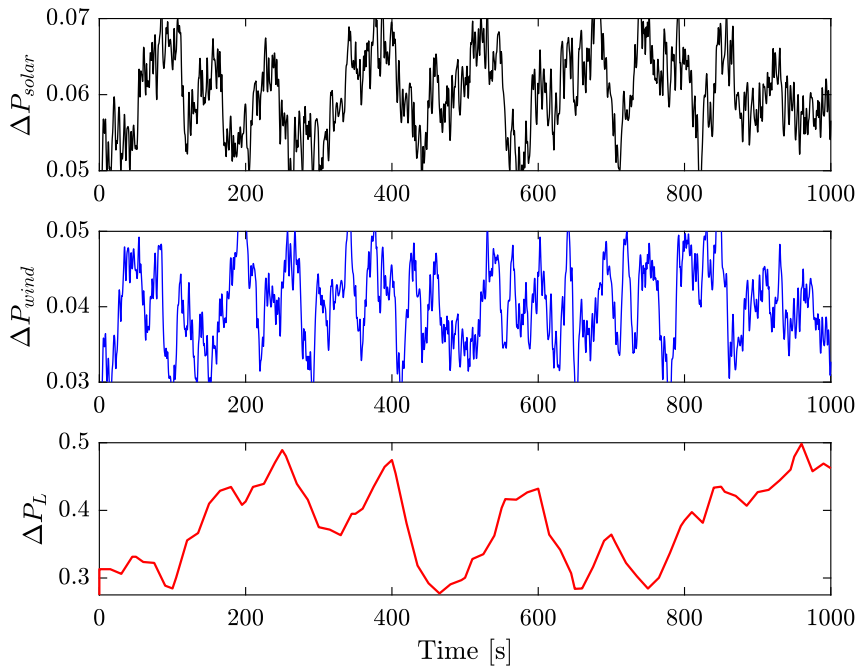
Figure 9: Power generated by wind and solar plants (top and middle), domestic load profile (bottom)
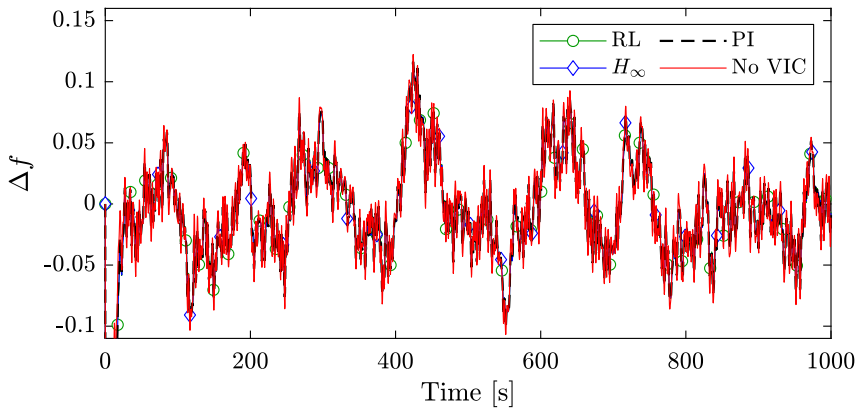

Figure 10: Frequency variation for the nominal case

Table 2: Calculated performance in Nominal Scenario

Metric	Controller	Nominal
RMSE	Proposed	0.0358
	H_∞	**0.0356**
	PI	0.0364
MAE	Proposed	**0.0287**
	H_∞	**0.0287**
	PI	0.0293
IAE	Proposed	**28.4136**
	H_∞	28.4461
	PI	28.7557

5.2 Deep reinforcement learning-based controller with advanced virtual inertia controller

5.2.1 Advanced virtual inertia controller

In [86], the novel virtual inertia controller is based on the nonlinear control strategy of the energy storage system presented in [47] and includes the positive (red) and negative feedback (blue) loops (see Fig. 11). The positive loop has the energy accumulation part E_s with defined charge limits $2E_{\max}$ and 0; the energy acceleration part is performed as the dead-band $\zeta(E_s)$ connected to controllable signal $\Delta\alpha$, which goes to the integral block with limits $\lambda_{\max}$ and $\lambda_{\min}$. The negative feedback loop has controllable signal $\Delta\beta$. Finally, both loops control the power output ΔP_{VI}. The control part of the VIC is performed by a multi-loop based controller with three output signals: two non-negative $\Delta\alpha$ and $\Delta\beta$; and one real $\Delta\gamma$, which controls the power inflow ΔP_e from the microgrid to the storage device. The proposed multi-loop DRL controller has several input signals: Δf, ΔP_M and ΔP_L. The Δf value depends on system inertia and its stability is important. Therefore, the DRL controller includes the Δf acceleration via differential block and the total deviation by discrete-time integration block.
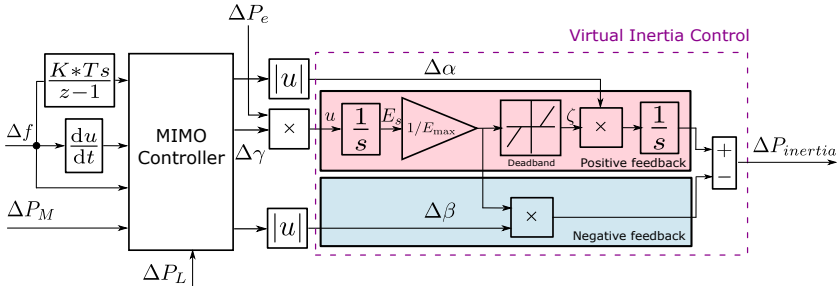

Figure 11: Schematic representation of the virtual inertia control strategy proposed in[86], including developed MIMO type of ANN-based controller/tuner

The considered microgrid model has an energy storage model (see Fig. 11), which provides power accumulation and (dis)charge dynamics. The dynamics of energy storage are

defined using the following equation:

$$\frac{\mathrm{d}E_s}{\mathrm{d}t} = u(t) = \Delta P_e \Delta \gamma,$$
$$E_s(0) = 0.5 E_{\max}, \; 0 \leq E_s \leq E_{\max}, \tag{21}$$

where ΔP_e is the power flowing into the storage and $\Delta \gamma$ is the control law. The dead-zone function is defined as:

$$\zeta(E_s) = \begin{cases} E_{\max}, & \frac{E_s}{E_{\max}} > \zeta(E_s)_{\max} \\ 0, & \zeta(E_s)_{\max} \geq \frac{E_s}{E_{\max}} \geq \zeta(E_s)_{\min} \\ -E_{\max}, & \frac{E_s}{E_{\max}} < \zeta(E_s)_{\min} \end{cases} \tag{22}$$

where the input is defined as $\frac{\mathrm{d}\lambda}{\mathrm{d}t} = \Delta \alpha \zeta(E_s)$, $\lambda(T) = 0$. The following part represents the ESS controllable power output that summarises all signals:

$$\Delta P_{VI} = \frac{\mathrm{d}\lambda}{\mathrm{d}t} - \Delta \beta \frac{E_s}{E_{\max}}. \tag{23}$$

5.2.2 Deep Reinforcement Learning-based Controller

The proposed deep reinforcement learning-based controller is a combination of two artificial neural networks: the actor $\mu(s \mid \theta_\mu)$ that preforms the frequency control and the critic $Q(s, a \mid \theta_Q)$ that selects the best actions of an actor (see Fig. 12). The training rules of the agent are defined by certain reinforcement learning (RL) policies (e.g. Dynamic programming, Monte Carlo, Q-learning, SARSA, DDPG, A3C, SAC) [81, 48, 91]. The major feature of RL training is a strategy based on the rule of trial and error. When the actor makes a correct action a_t according to the designed reward/punishment rules, it receives a reward r at each step of the action; otherwise, it receives a punishment $-r$. The training duration depends on the number of the epochs and learning rate α_μ actor and α_Q critic networks; after each epoch, the reward system calculates the average reward r_{avr}. If the average reward of the last epoch is significantly less than that of the previous, the algorithm initialises new weights of neural networks to try another combination, unless the accumulated reward reaches a target value r_{target}. We apply this architecture as an additional MIMO controller of the VIC scheme, which is illustrated in Fig. 11.
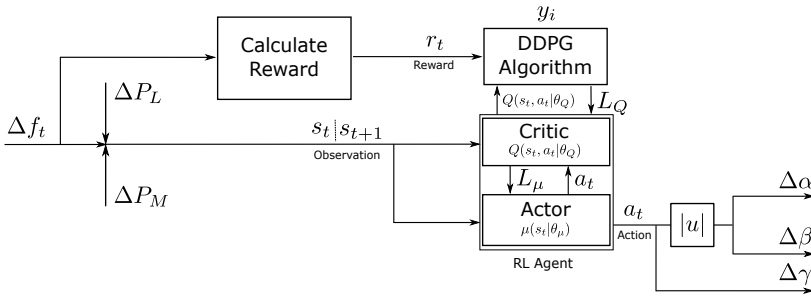

Figure 12: Schematic representation of implemented deep reinforcement learning control algorithm for energy storage model applied in [87] with consideration of dynamical signals for ESS $\Delta \alpha$, $\Delta \beta$ and control of incoming power by $\Delta \gamma$
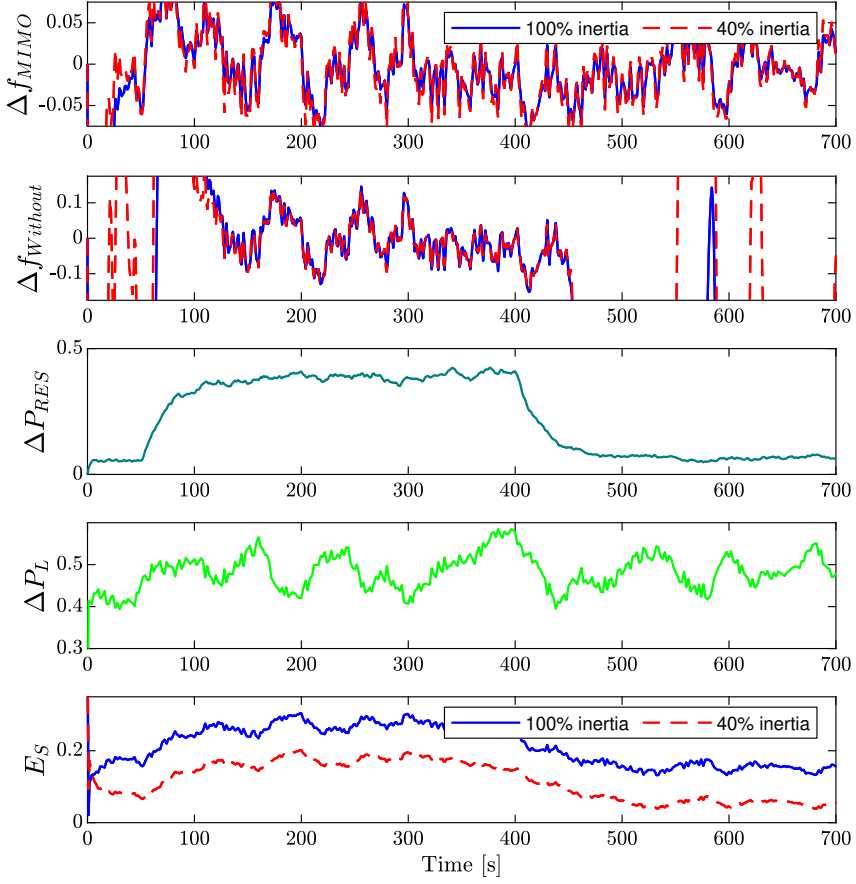
Figure 13: Scenario with simulated large amount of incoming power generated by park of wind turbines, organised to show (dis)charge changes of applied energy storage model

5.2.3 Numerical results

The simulated scenario allows the smooth connection of wind turbines at $t = 300$ s and disconnection at $t = 400$ s. It can be seen that the ESS charges faster after connection of the wind generated power and initiates discharge after disconnection of this power source as depicted in Fig. 13. As we can see before the connection of wind turbines, energy storage has a minimal state of charge. After disconnection of wind turbines, the energy storage provides active frequency support to the microgrid, helping to stabilise Δf in the case of the proposed controller; however, without the additional controller, the frequency keeps oscillating. In contrast to the previous scenario, the influence of low inertia on energy reservation is higher. Fig. 14 shows how the controller responds to the connection of a renewable source in events with excess power income, e.g. strong wind.

Table 3 shows the evaluation results of the proposed controller in a nominal scenario and a scenario with (dis)connecting renewable energy sources. Root mean square error (RMSE) and integral absolute error (IAE) statistical metrics are used to validate the performance. Both metrics confirm that the proposed VIC and DRL-based controller performs better than the standard one.
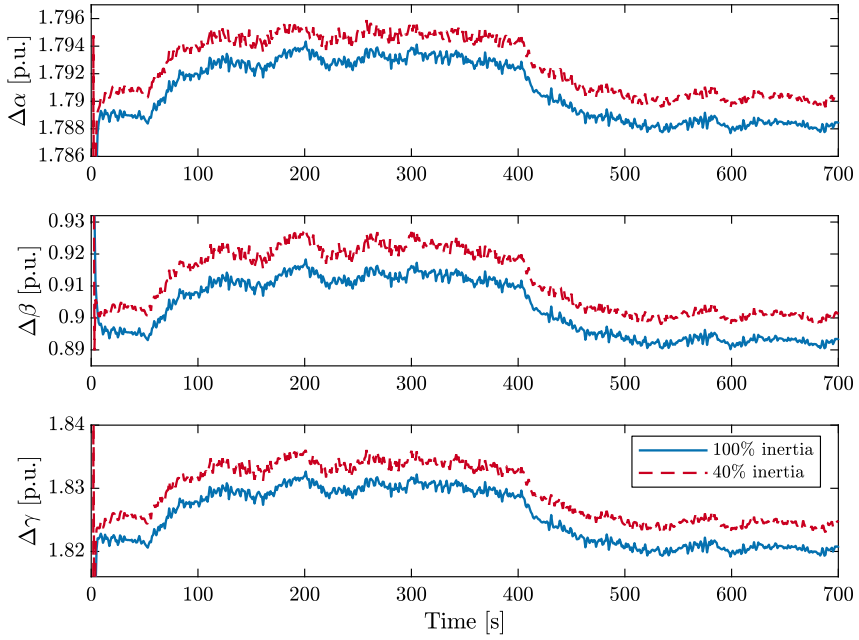
Figure 14: Provided control signals of ESS with dynamical coefficients Δα, Δβ and Δγ

Table 3: Performance of the Proposed Controller

Approach	With MIMO controller		Without MIMO controller	
Inertia	100%	40%	100%	40%
IAE	37.4870	38.7273	2040.9	1725.6
RMSE	0.0999	0.1027	3.0616	2.9267

5.3 Deep reinforcement learning-based VFOPID Controller

The key advantage of reinforcement learning is the direct interaction of an NN with a controllable environment and the application of nature-inspired reward-punishment-based learning, by which it can theoretically find the global minimum after the training procedure [45, 91]. Usually, reinforcement learning is applied in applications where it is necessary to train an NN for tasks in which the correct interaction of an autonomous system with the environment is critical [48, 91]. The general structure of the proposed controller is presented in Fig. 15. How to properly interact with VFOPID and MG using "double feedback", i.e. from the reward system and from the process directly, can be seen in the presented architecture of neural actor-critic studies, which are illustrated in detail in Fig. 7. The first operates in offline mode and the second in online mode, which can be summarised as follows:

- Reward or punish the agent by directing the measured error Δf into the block "reward system".

- Formulate the observation for the agent using the measured variables Δf and ΔP_E, which can be considered as the control error and pre-error, respectively.
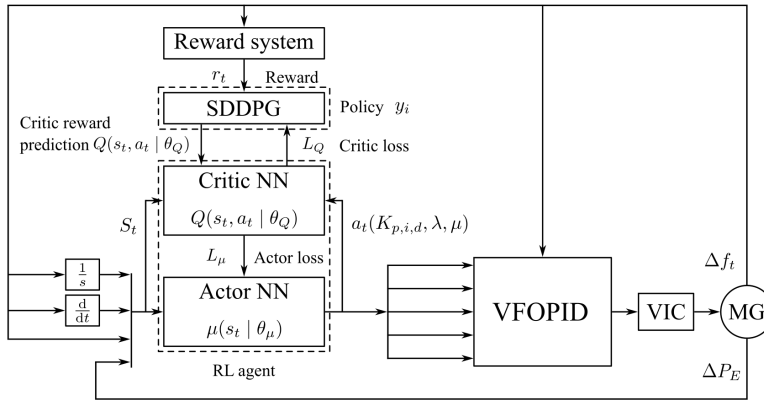

Figure 15: Scheme of the proposed DRL-based VFOPID for VIC in an MG proposed in [84]

5.3.1 VFOPID Controller

The original FOPID controller was first introduced in its parallel form described as [73]:

$$u(t) = K_p e(t) + K_i D^{-\lambda} e(t) + K_d D^{\mu} e(t), \qquad (24)$$

where D^{α} is the fractional differential ($\alpha > 0$) (or integral ($\alpha < 0$)) operator, μ and λ are the positive integration and differentiator orders, respectively, $u(t)$ is the control signal, $e(t)$ is the error signal, which is equal to $\Delta f(t)$, and K_p, K_i and K_d are the proportional, integration and derivative constant gains, respectively. The controller (24) has more tuning freedom that leads to making the plant stable under control and fulfilling intricate control performance requirements that are not in the scope of classical controllers. However, as this controller binds a control engineer to manipulate all of the features, the VFOPID controller with five variable parameters can significantly improve the performance of the controlled system due to its greater flexibility. Referring to Fig. 16, this means that it is possible not only to move continuously in the PID plane, instead of jumping between the fixed points, but also to explore for the desired controller values in a space inside the cube and between the eight vertices.
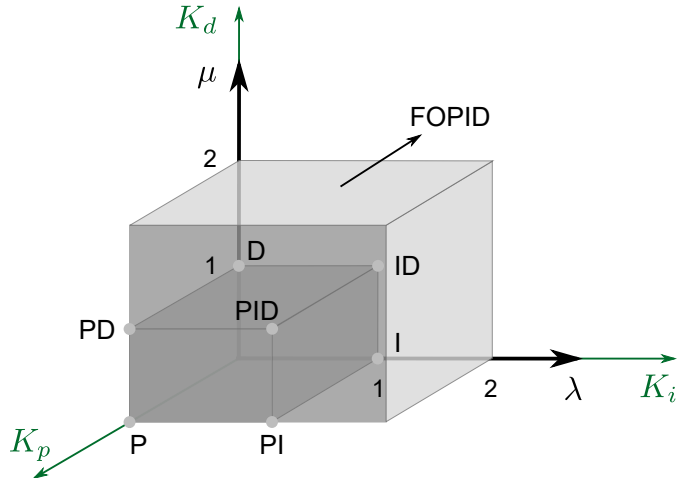

Figure 16: Relationships between the VFOPID controller and its classical variations

Integration and differential functions affect the steady-state process and the dynamic characteristics when both the associated gains and orders of each function vary:

- *Integral action:* This action is dependent on the gain K_i and the order λ. While K_i being oversized will make a system more unstable, its smaller size will force the system to diverge from its ideal dynamic performance. Additionally, for a small λ, the frequency band is wide, and the system is steady with rapid response and static error. In contrast, an oversized λ will degrade system stability with the increase in the overshoot, rise and settling times.

- *Derivative action:* The gain K_p has no effect on the steady-state error but can improve the dynamic characteristics. For a small K_p, the overshoot and settling time will increase, while for a large gain, the system noise may increase and the system performance will degrade. Finally, the order μ can improve response accuracy and

steady-state error when it is relatively small. However, by increasing μ, the over-shoot and settling time decrease and the closed-loop system stability will degrade.

It is necessary to keep both the gains and orders within a suitable range at any moment in order to maintain satisfactory control performance for a target system. To take advantage of this controller flexibility and promote controlled system performance simultaneously, the VFOPID needs to be considered as a target controller in the regulation process. The concept of VFOPID can be given as [50]:

$$u(t) = K_p(t)e(t) + K_i(t)D^{-\lambda(t)}e(t) + K_d(t)D^{\mu(t)}e(t), \tag{25}$$

where $D^{-\lambda(t)}$ and $D_t^{\mu(t)}$ are the variable order fractional integral and derivative, respectively. The implementation of (25) with a floating point requires a powerful computational device [70]. Therefore, we propose using a parallel connection of FOMCON library blocks from [94] and provide a real-time switch between fractional-order constants with a resolution of 0.1. In conjunction with switching objective functions and arising from an input signal derived from the NN tuner system, these functions attempt to switch between fractional integration and differentiator operators and implement a VFOPID controller according to the tuned parameters. By doing so, the fractional-order tuning and controller gains can be performed online.

5.3.2 ANN Tuning-based VFOPID Controller

We apply the multiple input multiple output type of NN as the tuner of the VFOPID controller. This hybrid algorithm combines the robustness of FOPID and the flexibility of ANN. The major task here for actor NN (i.e. tuner) is to search for an optimal combination of these coefficients considering MG disturbances. Fig. 17 illustrates the proposed architecture of the ANN-based tuner, where output neurons are five controller parameters: K_p, K_i, K_d, λ, and μ. Since the values of each output neuron should be non-negative, we designed the network to keep the outputs for the fractional orders and the gains of the controller within the range $[0, n/2]$, defined by the equation

$$y(t) = \frac{n}{1 + e^{-w|x(t)|}} - n/2. \tag{26}$$

In fact, the variable orders λ and μ are very sensitive to any changes and are computationally expensive. Therefore, we organised the change in the integral and derivative parameters with the discrete step of 0.1 using an automatic switch between predefined series of FOMCON blocks implemented in the MATLAB/Simulink environment, where each block has a frequency range $[0.001, 1000]$ and an approximation order 3. In the proposed method, tuning the VFOPID parameters depends on the magnitude of Δf being proportional to the power variation ΔP_E, which is driven to the NN tuner (see Fig. 17) and is then transformed by the *tansig* activation function illustrated in eq. 26 to give the tuned parameters as the given trajectories. The output value of every coefficient depends on the gain provided by every weight considered in the NN tuner.
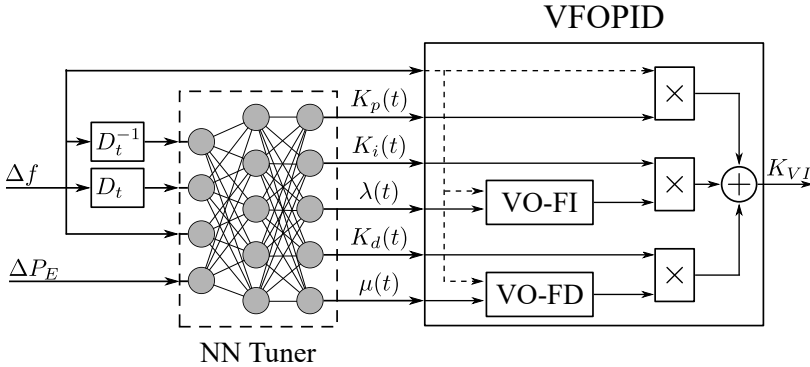

Figure 17: *Scheme of VFOPID controller with proposed NN-based tuner designed for [67], taking into consideration switchable FOMCON blocks of orders λ and μ defined by VO-FI and VO-FD, respectively*

5.3.3 Numerical results

We analysed the performance of the proposed algorithm (NN-based VFOPID controller) and compare it with FOPID and PID. The controllers were carefully tuned and the obtained parameters were $K_p = 50$, $K_i = 1.75$, $K_d = 2$, $\lambda = 1.25$ and $\mu = 1.75$ for the FOPID controller, and $K_p = 50.5$, $K_i = 5.85$ and $K_d = 5.5$ for the PID controller. These parameters were obtained by heuristic search, where we attempted to find the optimal parameters for the studied system and provide a fair comparison with the proposed method. In these experiments, we perform simulations for scenarios with (dis)connection of RESs for nominal (100%) and decreased inertia (40%). Fig. 18 shows the dynamics of renewable energy. To compare the algorithms, an additional (dis)connection-based scenario is carried out
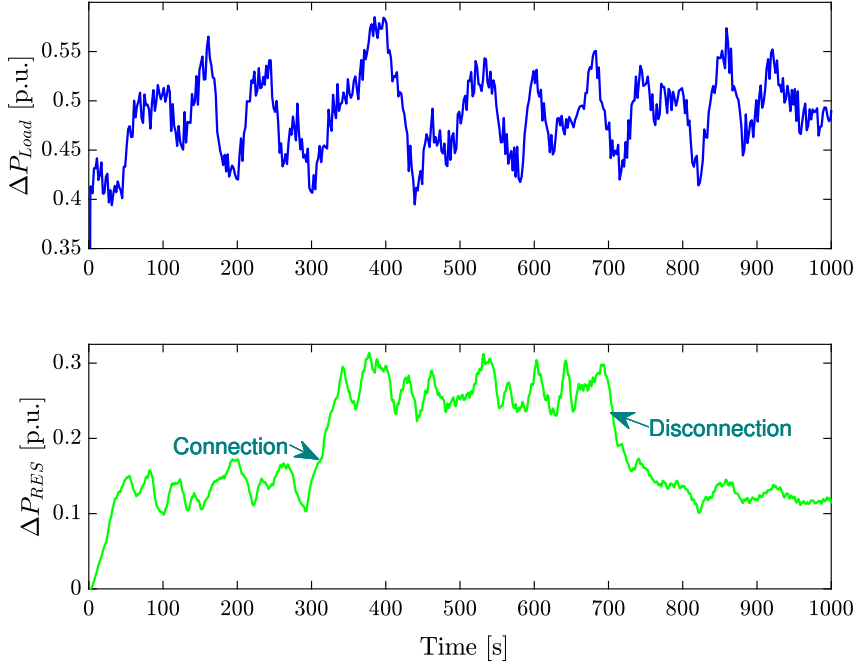
Figure 18: Data used in the considered scenarios with RESs connection and disconnection

and the performance is evaluated. We organise the smooth connection of solar panels and the disconnection of wind turbines in time steps 300 and 700 seconds, respectively (see Fig. 18). Fig. 19 depicts the significant influence of the wind turbine connection on the performance of all control algorithms. However, the proposed strategy shows the best performance in smooth transient behaviour. Performance indices are calculated in the steady-state conditions (i.e. from 100 to $1,000$ seconds) to eliminate the effect of high initial oscillations. All calculations are performed for three scenarios with both nominal (100%) and decreased inertia (40%). The results are summarised in Table 4. It can be seen that the proposed NN-based VFOPID controller has better results in all cases modelled.

Table 4: Comparison of different controllers using metrics

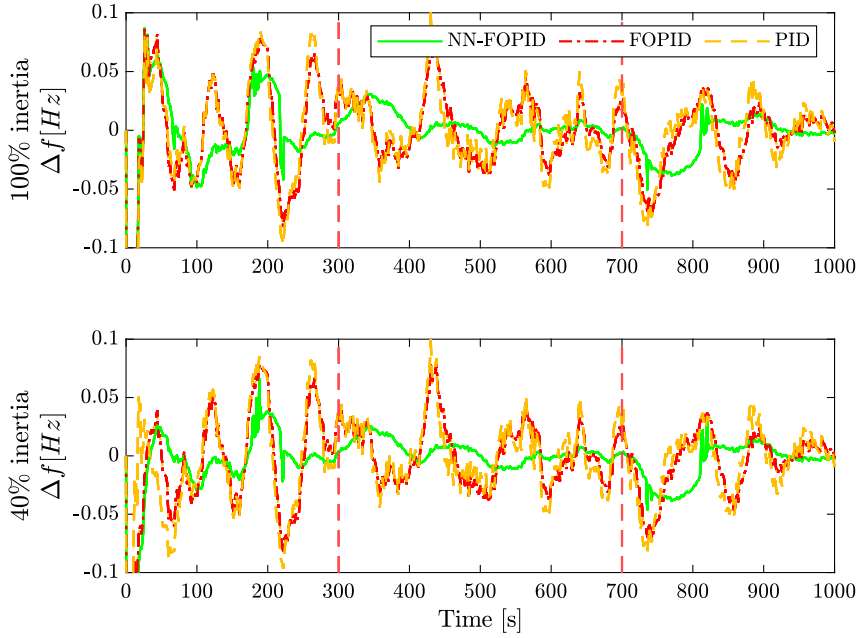| Metric	Inertia	Proposed	FOPID	PID
IAE	100%	**11.10**	21.23	22.81
	40%	**9.18**	21.26	22.91
RMSE	100%	**0.0175**	0.0300	0.0328
	40%	**0.0147**	0.0301	0.0330
MAE	100%	**0.0123**	0.0236	0.0253
	40%	**0.0102**	0.0235	0.0254

Figure 19: Frequency deviation. Performance results of the proposed (green), FOPID (purple with crosses) and PID (dashed orange) with different inertia. Red dashed lines indicate RES (dis)connection moments

5.3.4 Online Tuning of the VFOPID

Fig. 20 presents an example of the online optimisation (tuning) results of all five VFOPID controller parameters, where the output of the neural tuner is shown and the influence of renewable sources and decreased inertia on the VFOPID tuner is significant, where the integral knob K_i has a notable increase after connecting the wind turbine from 300 to 400 seconds and decreases after disconnecting the solar panels in 700 seconds. A similar change is seen in K_p, K_d, μ, and λ.
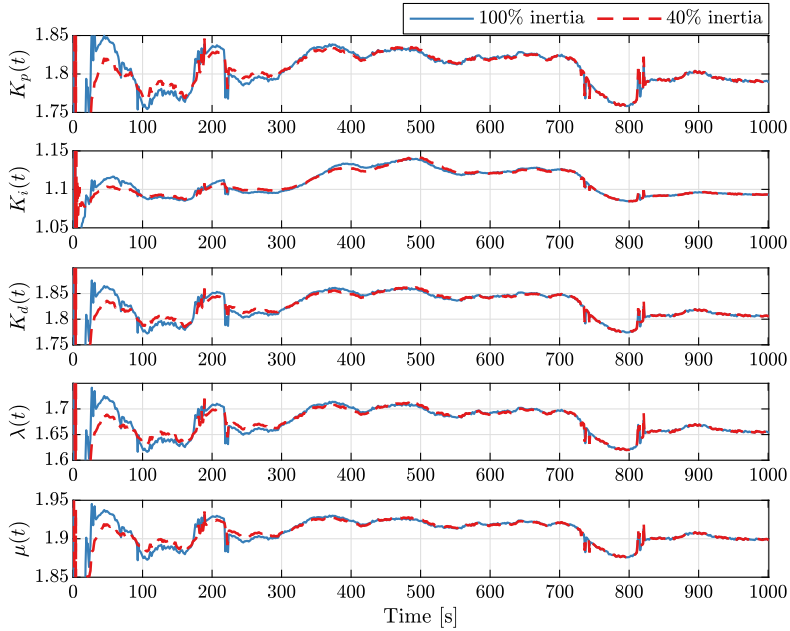
Figure 20: Evolution of VFOPID gains and orders for both nominal (100%, blue) and decreased inertia (40%, red)

5.4 Conclusion

Increasing the penetration of renewable source-based power creates multiple challenges related to the frequency stability of future microgrids. These problems can be solved by adding optimally designed control algorithms. Recent studies indicate that without extra virtual inertia support, the system is stable only for relatively high system inertia. Otherwise, the virtual inertia device is not able to provide satisfactory frequency stability with high penetration of renewable energy sources. Therefore, in this paper, we propose the reinforcement learning-based control of virtual inertia and compare this to H_∞ and PI controllers. In most situations, the proposed controller has shown slightly better results. More complex scenarios and detailed analyses will form a subject for future research. Reinforcement learning is a recently introduced optimisation strategy. Also, in this chapter, we presented the novel combination of an NN with a VFOPID controller, applied to robust virtual inertia emulation in isolated microgrids with high renewable energy penetration, where all tuned parameters (both gains and orders) of this controller were optimised by the semi-stochastic DRL algorithm with SDDPG policy. We proved that VFOPID controller fusion with an NN-based tuner is an effective strategy for virtual inertia emulation tasks. In contrast to other proposed algorithms, the NN-based VFOPID controller has active support from a machine-learning algorithm that makes it self-adaptive to MG disturbances, resulting in smooth deviation but a more significant influence from decreased inertia. It was shown that the proposed controller is capable of providing good response and frequency support in all simulated scenarios. The artificial neural network controller is faster but less flexible in many disturbances, including decreased inertia. Despite differences in the reinforcement learning algorithm applicable for training all tested controllers, every methodology requires design individual reward/punishment rules, including the sensitivity of the reward system to the measured error and limitations in rewarding.

6 Coordinated frequency control in hybrid microgrid

This chapter follows a discussion on the design of coordinated frequency control in the hybrid microgrid, which, in contrast to the isolated microgrid discussed in the previous chapter, has frequency control via energy management of the fuel cell and ultra capacitor. These two independent units require the integration of multi-agent control architecture to provide power balance.

6.1 General Structure of hybrid microgrid

Figure 21 presents the proposed hybrid MG, which encompasses power generation sub-systems consisting of a WTG, PV panels, an FC system and a DLC. These subsystems are interconnected in parallel to a shared AC bus that provides isolated loads. To examine this system, high-order mathematical models with nonlinear dynamics are required for each subsystem. However, in the majority of studies, first-order transfer functions are used to analsze all components, and the system simulations rely on simplified linear models. Although deploying nonlinear modelling techniques can greatly increase the accuracy and extract the full system potential, they come at a high computational complexity along with challenges in the LFC system analysis during the design process [102]. In this regard, we substitute the FC and DLC systems with their fractional-order models, which eliminate the disadvantage of the low accuracy while at the same time reducing the model complexity.
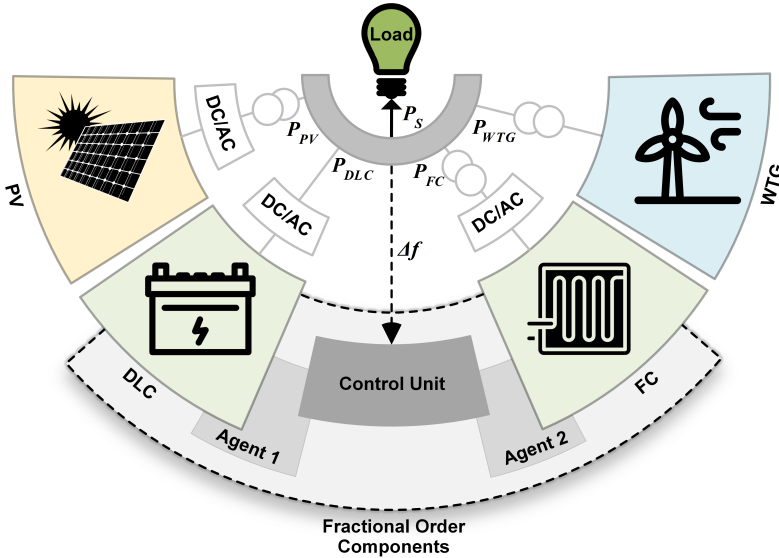

Figure 21: General structure of the hybrid microgrid

Effective control of components is necessary for the stable operation of the hybrid microgrid (HMG) system with different power generations. This can be achieved by regulating the fluctuation in the frequency profile Δf, which can be expressed using the power balance $\Delta P = P_{Net} - P_{Load}$ as

$$\Delta f = \frac{\Delta P}{K_{HMG}}, \tag{27}$$

where P_{Net} and P_{Load} are the net and load power, respectively, and K_{HMG} is the system frequency constant of HMG. The dynamical model of frequency variation according to per

unit power deviation can be given by

$$\frac{\Delta f}{\Delta P} = \frac{1}{sM + D},$$ (28)

where M and D are the inertia and damping constants of HMG, respectively. Using a coordinated control strategy between FC and DLC (see Fig. 23), the FDC problem can be solved with enhanced power quality. In this regard, DLC and FC systems act as backup systems and compensate for high and low frequency deviation, respectively. A high-pass filter (HPF) here can reduce the charging and discharging of DLC in long-term operation. Also, as seen from Fig. 23, the net power generation is comprised by the following equation:

$$P_{Net} = P_{WTG} + P_{PV} + P_{FC} \pm P_{DLC}.$$ (29)

Table 5: Parameters of the isolated HMG

RESs	FC and DLC	Frequency Deviation Model
$T_{WTG} = 1.5$ s	$T_{FC} = 0.26$ s	$M = 0.4$
$T_{PV} = 1.8$ s	$T_{DLC} = 0.01$ s	$D = 0.03$

6.1.1 WTG Model

Wind speed has a direct effect on the output power of WTGs, as mechanical power in the turbine is determined by the formula

$$P_{WTG} = \frac{1}{2}\rho A v^3 C_p(\lambda, \theta),$$ (30)

where ρ is the density of air, A is the area swept by the blades, and v is the wind velocity. Also, C_P is the rotor efficiency, which is a function of tip speed ratio λ and pitch angle θ, as

$$C_p(\lambda, \theta) = 0.5(116\hat{\lambda}^{-1} - 0.4\theta - 5)e^{-21\hat{\lambda}^{-1}} + 0.0068\lambda,$$ (31)

with $\hat{\lambda}^{-1} = (\lambda + 0.08\theta) - 0.035(\theta^3 + 1)$. As depicted in Fig. 22a, the $C_{P,max}$ can be acquired for a given direction of the blades and when λ is in its special value, which directly depends on the aerodynamic structure of the turbine. Depending on wind velocity v, the rotor speed can keep λ at its optimal level, which means that the most energy from the wind can be used. Also, Fig. 22b illustrates the variation of P_{WTG} according to v. It remains constant by using the pitch angle control system to prevent excessive rotor speed and preserve the equipment when v increases to the rated wind velocity. When v is smaller and greater than the cut out and on wind speeds, respectively, i.e. $14 < v < 25$ m/s, P_{WTG} takes its constant maximum value and is zero for $v < 4$ m/s. However, for $v > 25$ m/s, the wind turbine stops operation. By defining T_{WTG} as a time constant, the dynamical model of WTG in the frequency domain can be given by

$$\frac{\Delta P_{WTG}}{\Delta P_{wind}} = \frac{1}{sT_{WTG} + 1},$$ (32)

where ΔP_{wind} and ΔP_{WTG} are the variations of the mechanical power and output power of WTG, respectively.
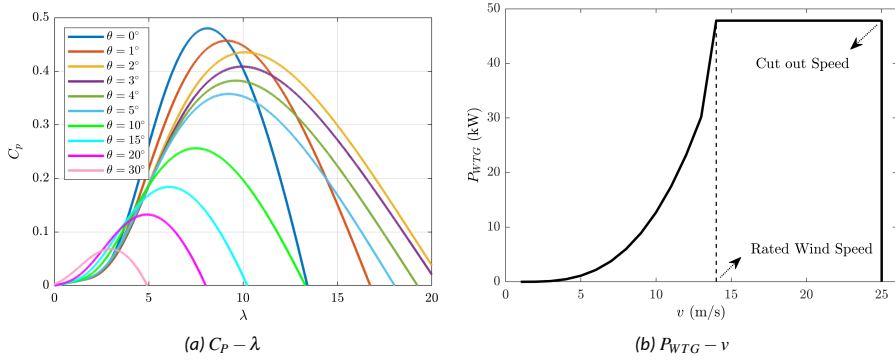
(a) $C_P - \lambda$ *(b) $P_{WTG} - v$*

Figure 22: Variation of C_P and P_{WTG} according to the λ and v.

6.1.2 PV Model

One of the most promising, flexible and environmentally friendly power sources is the PV system, which consists of photovoltaic panels connected together in series and parallel structures. This system converts solar radiation to electrical energy based on solar cell temperature and array area, and its output power P_{PV} can be expressed as follows in terms of conversion efficiency η, PV array area S, solar irradiation φ and ambient temperature T_a:

$$P_{PV} = \eta S \varphi (1 - 0.005(T_a + 25)). \tag{33}$$

The variations of solar irradiation to output power of PV system can be described as

$$\frac{\Delta P_{PV}}{\Delta \varphi} = \frac{1}{sT_{PV} + 1}, \tag{34}$$

where T_{PV} is the PV time constant.

6.1.3 FC Model

The FC technology, which has been considered as a high-efficiency power generation device, can convert chemical energy into electrical energy using hydrogen and oxygen. This static device has slow dynamics in its fuel supply sections, such as pumps and valves, resulting in slower power output. The dynamic response of the FC system can be given by introducing the first-order time delay transfer function and T_{FC} as the FC time constant

$$G_{FC} = \frac{1}{sT_{FC} + 1}. \tag{35}$$

6.1.4 DLC Model

The General Electric Company first patented electro-chemical capacitors in 1957, consisting of porous electrodes using the DLC mechanism of charging [100]. Nowadays, this technology plays a key part in fulfilling the demands of possible applications, including memory back-up, electric vehicles, power quality and RESs. The transfer function of this highly efficient device with a fast load frequency can be represented by the first-order lag equation

$$G_{DLC} = \frac{1}{sT_{DLC} + 1}, \tag{36}$$

where T_{DLC} is a time constant.

6.2 Coordinated ANN-based PI controller

In the proposed coordinated control solution, each ANN performs online tuning of each PI controller simultaneously. At time step t and after each system disturbance, the given solution provides a reasonable change in the tuning of two parameters $K_P(t)$ and $K_I(t)$ of the controller, stated as

$$K_{PI}(s) = K_P(t) + \frac{K_I(t)}{s}. \tag{37}$$

The tuning of the differential part $K_D(t)$ is avoided because it is very sensitive to disturbances and difficult to change without losing tone stability. To organise the effective and fast training of two independent agents, we designed the multi-agent SRL optimisation method with individual positive weight w ranges $[0, 5]$ and $[0, 10]$ for agents 1 and 2, respectively. We also propose an ANN with a modified `tansig` activation function defined as

$$y = \frac{m}{1 + e^{-0.1w|x|}} - n, \tag{38}$$

where the absolute value of x is used to avoid negative output, provide a sufficient range of coefficients for the PI controller and make the tuning process more robust. For effective coordination purposes of the two subsystems, DLC and FC, the parameters m and n are considered to be 200 and 100 for the DLC element, and 10 times larger for the FC subsystem. We avoid connecting the HPF to the ANN-based tuner since it has a negative impact on the stability of the tuner and confuses the reward system. Therefore, the input signal for both agents is the same.
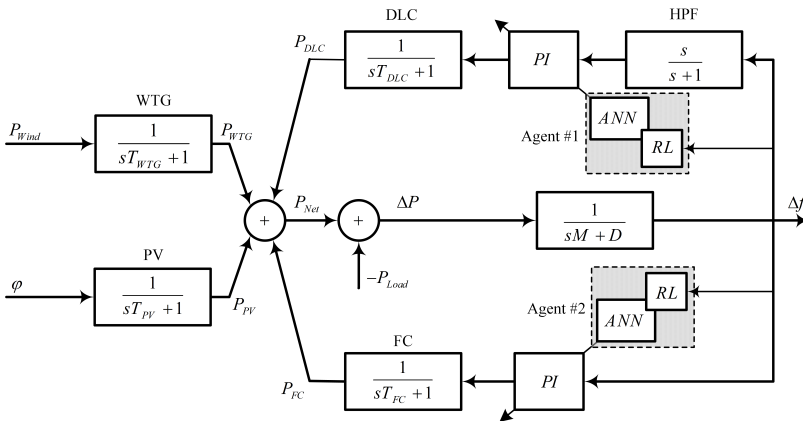

Figure 23: Proposed in [89]: coordinated multi-agent FDC scheme with NN-PI controller

6.2.1 Numerical Results

To develop the proposed multi-agent management strategy and to investigate the coordinated NN-PI controller performance, the given control scheme (see Fig. 23) has been implemented using MATLAB/Simulink with model parameters given in Table 5. Figure 24 (top and middle plots) shows the output power of WTG and PV systems modelled as day dynamics with period 24 hours; while the bottom plot depicts the load demand. Figure 34 shows how the FDC problem can be appropriately solved using the proposed controller, in which all hybrid power generation is better compensated by considering the effects of system frequency variation as compared to the usual PI controllers and when the system

is free of control. In terms of transient response, NN-PI has a smaller overshoot with a shorter settling time, indicating a faster transient time. In terms of steady state, the proposed approach has small steady-state errors when compared to the usual controller (see Table 6).
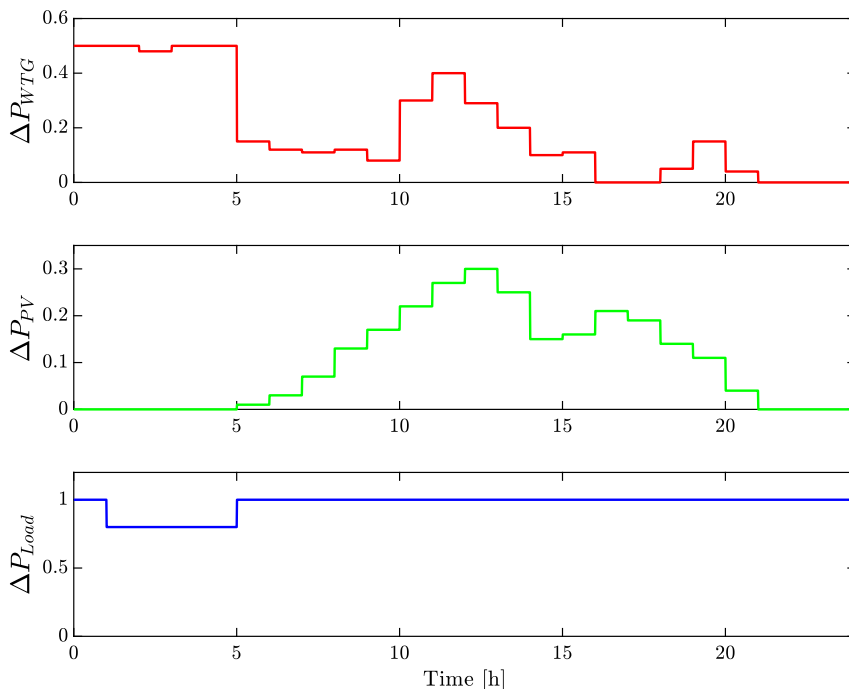

Figure 24: Output powers of WTG, PV and load variations

According to the results, we can see the insignificant influence of RESs disturbances on the proposed NN-PI controller. In contrast, the usual PI controllers show less flexibility in the distributive process. Figure 25 illustrates the online tuned parameters of two controllers. Here, we can see how the ANN adjusts the parameters of both controllers to optimal values after each time step. For the usual PI controllers, the parameters are fixed at $K_{p,DLC} = 30$, $K_{i,DLC} = 0.1$, and $K_{p,FC} = 150$, $K_{i,FC} = 1000$. Since the tuning process depends on measured error dynamics, each ANN provides the dynamics of tuning in a similar form. While in agent 1, the range of coefficients is similar; in agent 2, it is better to keep the coefficients different.

Table 6: Performance Results

Controller	IAE	Root MSE	MSE
Coordinated NN-PI	**0.0027**	**0.0004**	**0.0000016**
Coordinated PI	0.005	0.0011	0.000012

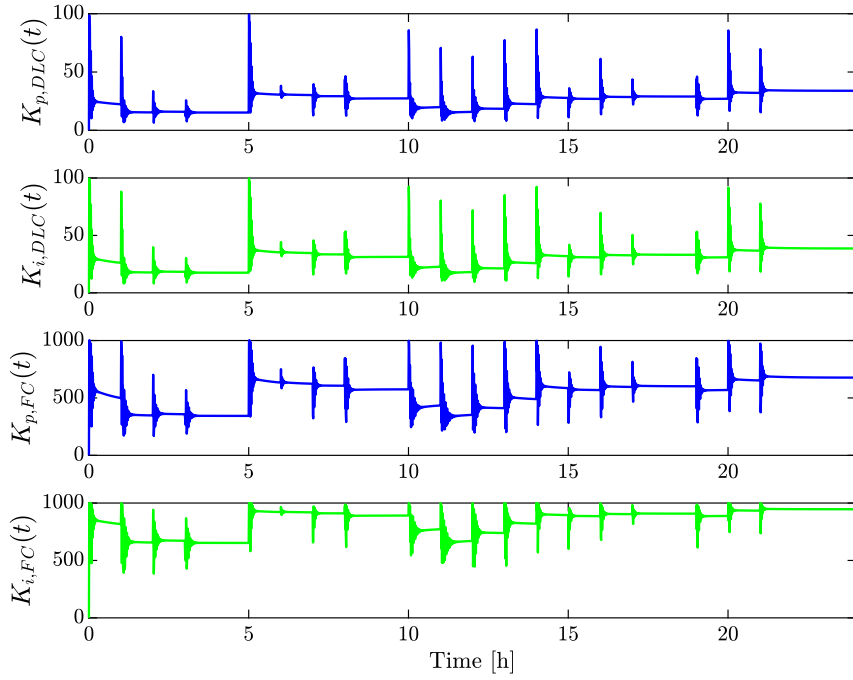IAE: Integral Absolute Error, MSE: Mean Square Error

54

Figure 25: Online tuning of PI controllers DLC (top) and FC (bottom)
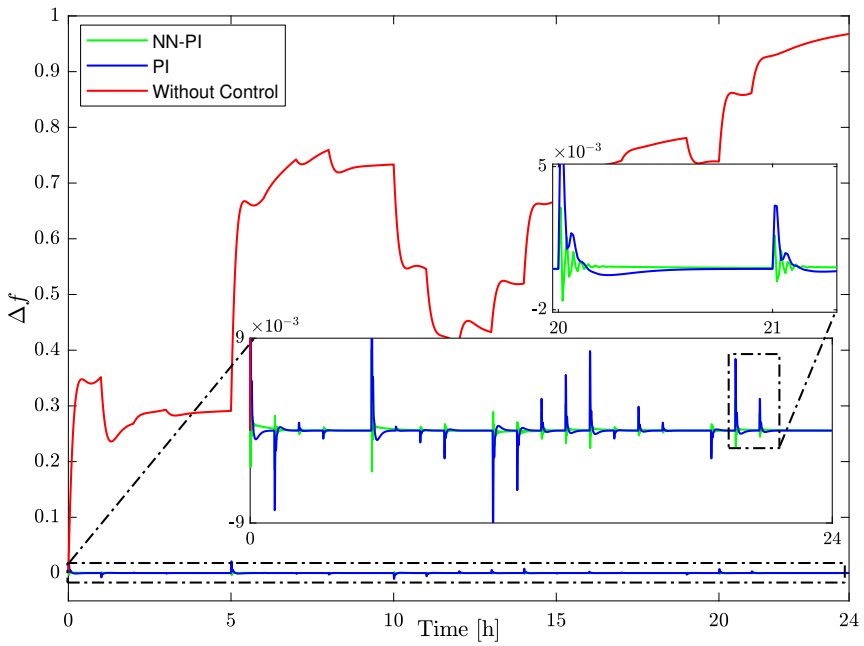

Figure 26: Control results of proposed control approach and PI controllers for DLC and FC

6.3 Coordinated ANN-based VFOPID control

The artificial neural network-based online tuner of (FO)PID extends the robustness of the FO(PID) controller and can be a direct competitor of the fuzzy logic-based (FO)PID, which can be optimised by a countless number of machine learning methods. The main idea of an online tuner is the design of architecture that uses a control error signal to change the values of (FO)PID. Machine learning helps to define the change in regard to which co-efficients are more significant; such an intelligent approach can automatically adapt the controller to any process. It is difficult say which architecture of neural network-based tuner is right, for example, the optimal number of hidden neurons for each process can be individual and an NN can take into consideration additional inputs that define control error, including setpoints, disturbances, etc. We suggest this problem can be solved by genetic algorithms such as the NEAT (neuroevolution of augmenting topologies) algorithm that changes the topology of a neural network. The ANN-based tuner must abide by the idea of a PID controller; in other words, the neural network-based tuner must include integrator D_t^{-1} (see Fig. 27), but differential part D_t is optional since is sensitive to disturbances in a control loop and can destabilise the tune of all (FO)PID coefficients. Moreover, negative values of PID ($K_p(t)$, $K_i(t)$, $K_d(t)$) are not acceptable, therefore activation functions in a neural network must be designed with this approach to avoid negative output. Nevertheless, variable $K_d(t)$ is very sensitive to any rapid change or controller error and cannot be stable for all processes, hence for the majority of processes, the variable FO(PI) version is more applicable. In the case of variable fractional order coefficients ($\lambda(t)$, $\mu(t)$), we have to be sure about the exact range and resolution of output values because they are computationally very complex in real-time control.
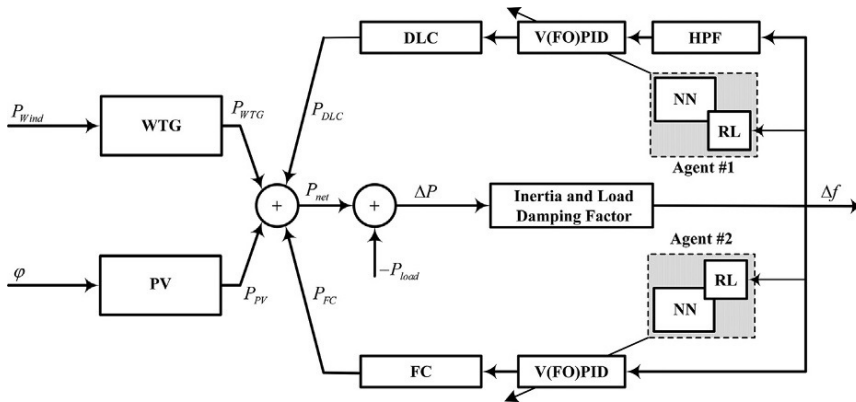

Figure 27: Proposed for [68]: coordinated multi-agent FDC scheme with V(FO)PID controller

6.3.1 Numerical results

According to the results, we can see an insignificant influence of WTG and PV disturbances on the proposed combination of NN and variable PID-type controller (see Fig. 28). In contrast, the traditional controllers show less flexibility in this case. In terms of transient response, the NN-based V(FO)PID controller has a smaller overshoot with a shorter settling time, thus indicating a faster transient time and hence better reference tracking performance. In terms of steady state, the NN-based controllers have small steady-state errors and reduce the values of the integral absolute error (IAE), mean square error (MSE) and root MSE (RMSE) when compared to the traditional PID type controller (see Table 7). As

Performance metric	NN-VFOPID	NN-PID	FOPID	PID
IAE	**0.003324**	0.004559	0.005774	0.006403
RMSE	**0.0001804**	0.0002482	0.000322	0.000323
MSE	**$3.2528*10^{-8}$**	$6.1618*10^{-8}$	$1.0633*10^{-7}$	$1.0416*10^{-7}$

seen, the proposed combination of NN and V(FO)PID produces the best results. One can see that when NN is combined with V(FO)PID, the reactions to the system disturbances are reduced. Figs. 30a and 30b illustrate the parameters of the online tuned V(FO)PID controller for both agents 1 and 2. Here, we can see how the NN adjusts the parameters of both controllers to optimal values after each step in the signal.

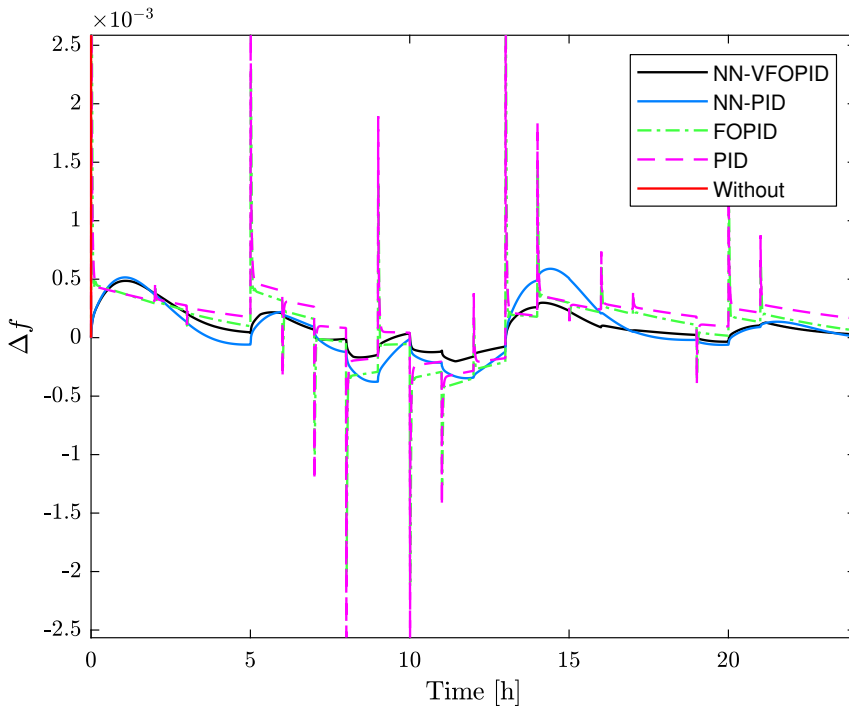*Figure 28: Performance of proposed coordinated NN-VFOPID control of hybrid microgrid with comparison of tested algorithms*
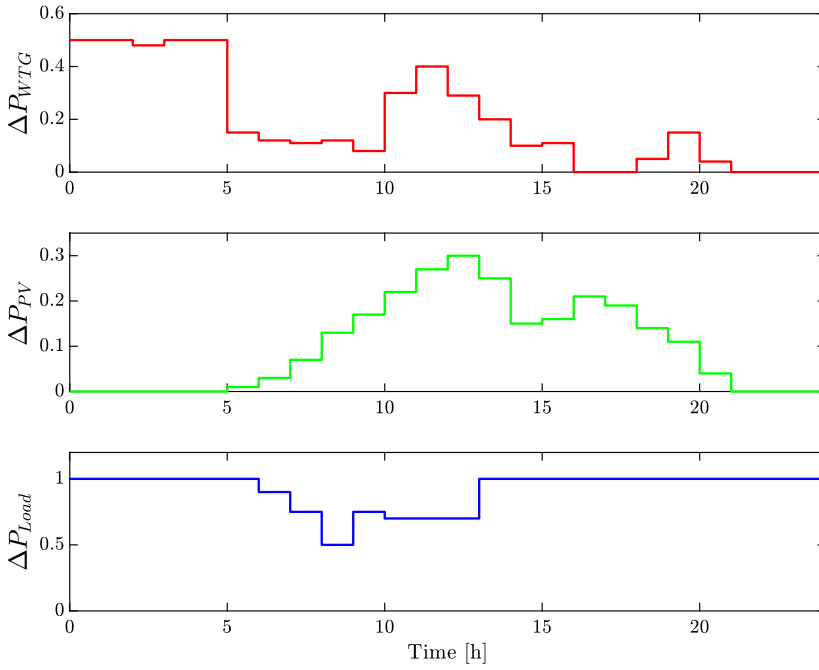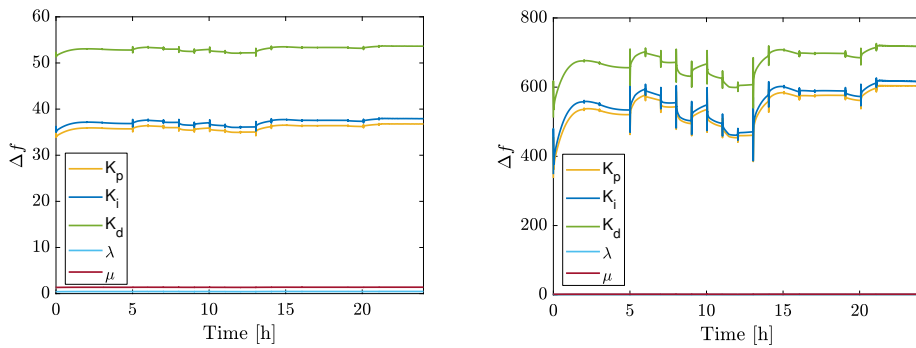
Figure 29: Modelled scenarios of renewable sources and loads in hybrid microgrid

(a) Online tuning of Agent 1 for FC control

(b) Online tuning of Agent 2 for DLC control

Figure 30: Online tuning of VFOPID controllers include dynamics of orders λ and μ, and gains K_p, K_i, K_d

6.4 Coordinated ANN-based control

Utilising a coordinated control strategy that combines FC and DLC systems, the LFC problem can be effectively resolved while also improving power quality, as depicted in Fig. 31. In this approach, the fractional order DLC and FC systems serve as backup systems and compensate for high and low frequency deviations, respectively. Also, to minimise the charging and discharging of FDLC during long-term operation, a high-pass filter (HPF) is employed. Furthermore, to maintain the stable operation of the autonomous isolated MG system, effective control of the supply power is necessary since the output power of various power generation components can fluctuate under certain conditions. The control strategy in this system is based on the power balance error ΔP, which is calculated as the difference between the power supply P_{Net} and the power demand P_{Load}, given as $\Delta P = P_{Net} - P_{Load}$. As power generation varies, the frequency fluctuates, and this frequency deviation Δf is calculated using

$$\Delta f = \frac{\Delta P}{K},$$ (39)

where K is the system frequency characteristic constant of the system. However, due to delays in the frequency characteristics, the above equation is modified to

$$\Delta f = \frac{\Delta P}{K(sT_f + 1)} = \frac{\Delta P}{sM + D},$$ (40)

which takes into account the frequency characteristic time constant T_f as well as the load damping constant D and the inertia constant M. The net power generation is computed using $P_{Net} = P_{WTG} + P_{PV} + P_{FFC} \pm P_{FDLC}$, which involves various contributing factors. For control purposes, two ANN-based controls are employed for two fractional subsystems associated with a common stochastic RL-based training unit. This control framework and its strategy, which is integrated into the control unit, will be further investigated in the following two subsections.
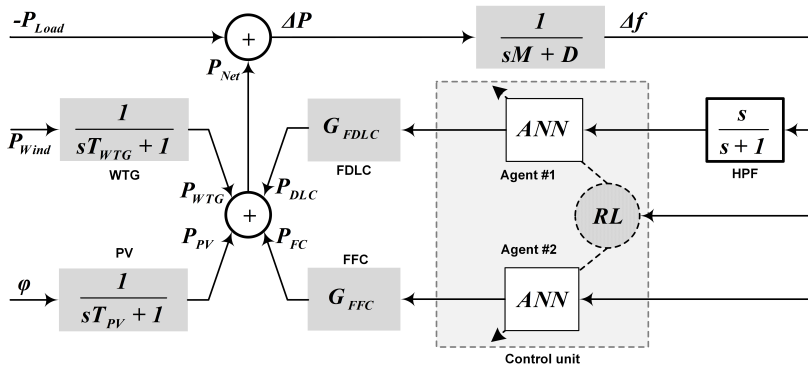
Figure 31: Coordinated LFC strategy with the multi-agent ANN strategy proposed in [88]

6.4.1 ANN-based controller

The proposed ANN-based controller, which is similar to the PID-type controllers, is illustrated in Fig. 32. The controller takes frequency deviations as an error signal in three different forms: original Δf, differentiated $d\Delta f/dt$, and integrated $\Delta f/s$, and inputs them into neurons. In this architecture, multiple ANN weights of actor systems are defined as

$w_{\mu1,i}$ and $w_{\mu2,i}$ for hidden and output layers, respectively. The hidden layer consists of three neurons, while the output layer has one neuron that summarises all signals into output provided for the inputs of fractional DLC and fractional FC systems. Every artificial neuron uses an adjustable `tansig` activation function that provides bipolar output and is suitable for many control applications. The performance of the proposed controller depends on the magnitude of the weights in each artificial axon and the constants defined in the applied activation function as

$$y(t) = \frac{n_i}{1 + e^{-w_i x(t)}} - \frac{n_i}{2}, \tag{41}$$

where weight w defines sensitivity to disturbances, constant n defines limits for output signal in every neuron and i is agent number
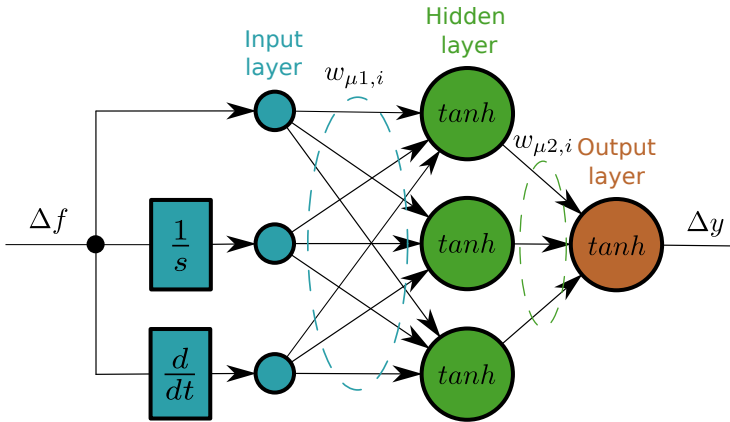

Figure 32: Proposed artificial neural network-based controller

6.4.2 Numerical Results

The proposed controller was found to be effective in addressing the LFC problem, as shown in Fig. 34 (top plot), where it outperformed the FOPID controller. In terms of transient response, the ANN-based controller demonstrated faster response times with smaller overshoots and shorter settling times in a fractional order environment. The proposed approach also exhibited smaller steady-state errors compared to the traditional controller, as detailed in Table 8. Our proposed algorithm yields lower values of integral absolute error (IAE), root mean square error (RMSE) and mean absolute error (MAE) when compared to the FOPID controller, as evidenced by the results. It is worth noting that for two FOPID controllers, the parameters are set to fixed values as $K_{p,DLC} = 100$, $K_{i,DLC} = 50$, $K_{d,DLC} = 0.5$, $\lambda_{DLC} = 1.35$, $\mu_{DLC} = 1.2$ and $K_{p,FC} = 1000$, $K_{i,FC} = 200$, $K_{d,FC} = 0.5$, $\lambda_{FC} = 1.2$, $\mu_{FC} = 1.25$. According to the results, we can see the insignificant influence of renewable energy source disturbances on the proposed ANN controller. In contrast, the usual FOPID controllers show less flexibility in the distributive process. Here, we can see how the ANN is robust and adjusts the parameters of both controllers to optimal values after each time step.

   To verify the effectiveness of the fractional order modelling strategy and its impact on the LFC enhancement compared to the classical integer order models, we repeated the experiment for the considered MG system for the case of integer order FC and DLC models with the same parameters. It can be seen in Fig. 34 (bottom plot) that the proposed

method using the same trained process for the fractional order case again shows better performance than the FOPID controller. One may see that this resulted in a slightly degraded performance when compared to the fractional order case above. Nevertheless, in both cases, the proposed controller was able to keep the frequency within the limits.

Table 8: Numerical comparison between ANN-based controller and FOPID

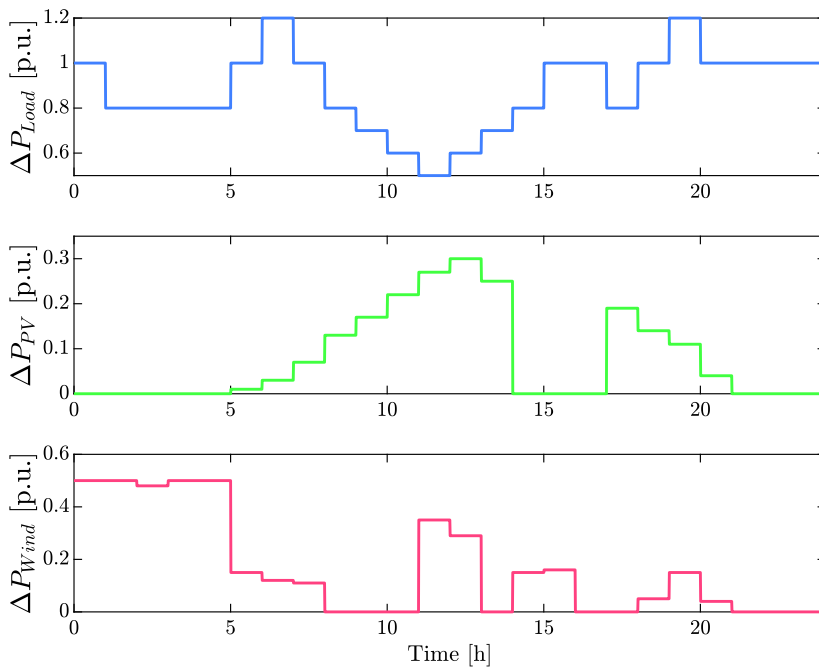| Controller	IAE	RMSE	MAE
Proposed	**0.0221**	**1.7678^{-5}**	**9.2214^{-6}**
FOPID	0.3125	1.7269^{-4}	1.3014^{-4}

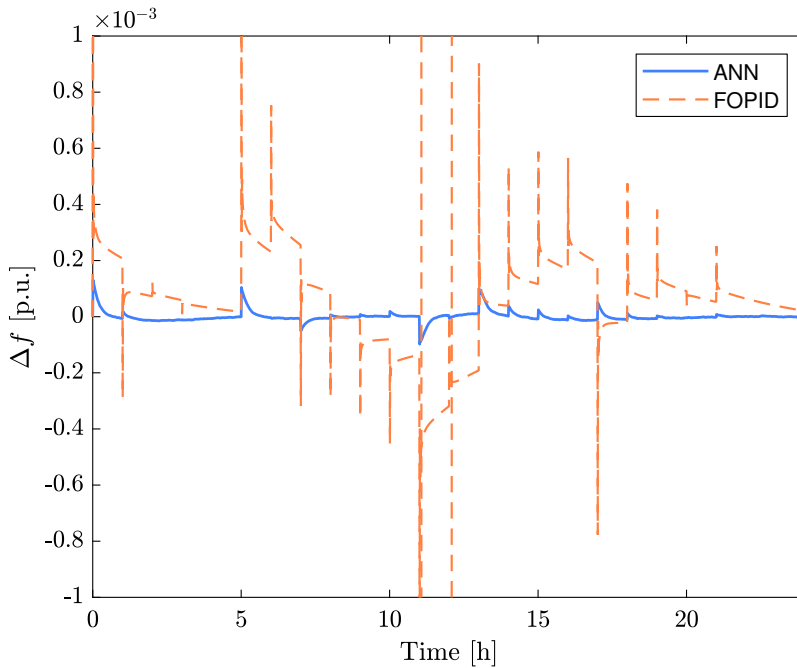*Figure 33: Output power changes of renewable energy sources and load*

Figure 34: Control performance of tested algorithms with fractional order system

6.5 Conclusion

In this chapter, we proposed ANN trained control methods using the SRL algorithm to solve the FDC problem of the given isolated HMG. The results demonstrate that the proposed combination of PI controller and neural tuner and automatic tuning technique captures all the remarkable advantages of the controller, including robustness improvement and disturbance rejection, and gives rise to a generation of control laws that improve both transient and steady state cases. Furthermore, a neural network as a dependent controller shows good results with the fractional order version of a hybrid microgrid. Two ANN controllers were deployed for each component in a multi-agent framework using a multi-agent SRL optimisation technique, which, in addition, were compared with two optimally tuned FOPID controllers. Test results on an isolated MG with fractional components validated the effectiveness of the proposed coordinated LFC strategy. As illustrated in this study, it was observed that while the proposed algorithm outperforms the FOPID controller in both systems with integer and fractional model of components, it exhibits an enhanced LFC system with greater stability in the MG system with fractional order subsystems. In addition, this chapter presented two NN-V(FO)PID controllers applied to the LFC problem in isolated HMG with high RES penetration, where the SRL algorithm optimised all tuned parameters (both gains and orders) of this controller. We demonstrated that combining a V(FO)PID controller with an NN-based tuner is an effective strategy for FDC tasks. Unlike traditional PID-type controllers, the NN-based V(FO)PID controller has active support from an ML algorithm, allowing it to self-adapt to HMG disturbances and produce smooth frequency deviation. The results demonstrate that the proposed combination of PID-type controllers and an NN tuner captures all the remarkable advantages of the controller and gives rise to a generation of control laws that improve both transient and steady-state errors.

7 Conclusion and Future Plans

This doctoral dissertation developed an artificial neural network-based controller with a training system provided by the deep reinforcement learning approach, which was applied in a simulated model of an isolated microgrid with high penetration of renewable energy sources. In this work, the author tried develop new methods that can increase efficiency and provide adaptive virtual inertia emulation. To provide effective training for an RL agent, the new reward system with designed frequency control rules was proposed. In the numerical results and control quality criteria presented in [87], we can see the proposed methods are competitive, on the one hand, to the classical control methods, including PID, H_∞, CDM, and on another hand to the advanced control methods, including MPC, FLC. Work [86] attempted to combine the control scheme of energy storage proposed in [47] with the designed MIMO neural network-based controller. The result of this work was increased performance and flexibility in the frequency support of an isolated microgrid. In publication [89], we tried to combine the fractional order PID controller with the developed artificial neural network-based tuner, which provided online tuning of FOPID knobs. With integrated reinforcement learning, this hybrid method showed very good results in frequency management.

These developments have a lot of potential for future development and research towards practical application. To make this more feasible, we plan and design a variable fractional order with direct order calculation instead of switching between predefined blocks. Some ideas in some publications were not fully completed; for example, the SD-DPG policy needs to be deeply integrated with stochastic reinforcement learning and be combined with a stability analysis of proposed control methods in order to design constrained neural network that can guarantee stable performance in power management. Moreover, we plan to implement an ANN as a controller for a real DC/AC inverter in order to provide robust and adaptive voltage control with respect to the measured current. Additionally, we plan to extend the designed system in a simulated multi-microgrid architecture, where a multi-agent control architecture is used to provide frequency stability in each area of the interconnected microgrid.

List of Figures

List of Tables

References

[1] H. Abubakr, T. H. Mohamed, M. M. Hussein, J. M. Guerrero, and G. Agundis-Tinajero. Adaptive frequency regulation strategy in multi-area microgrids including renewable energy and electric vehicles supported by virtual inertia. *International Journal of Electrical Power & Energy Systems*, 129:106814, 2021.

[2] H. Ali, G. Magdy, B. Li, G. Shabib, A. A. Elbaset, D. Xu, and Y. Mitani. A new frequency control strategy in an islanded microgrid using virtual inertia control-based coefficient diagram method. *IEEE Access*, 7:16979–16990, 2019.

[3] H. Ali, G. Magdy, and D. Xu. A new optimal robust controller for frequency stability of interconnected hybrid microgrids considering non-inertia sources and uncertainties. *International Journal of Electrical Power & Energy Systems*, page 106651, 2021.

[4] R. Ali, T. H. Mohamed, Y. S. Qudaih, and Y. Mitani. A new load frequency control approach in an isolated small power systems using coefficient diagram method. *International Journal of Electrical Power & Energy Systems*, 56:110–116, 2014.

[5] D. Bazargan, S. Filizadeh, and A. Gole. Stability analysis of converter-connected battery energy storage systems in the grid. In *2015 IEEE Power & Energy Society General Meeting*, volume 5, pages 1204–1212, 2014.

[6] H.-P. Beck and R. Hesse. Virtual synchronous machine. In *International Conference on Electrical Power Quality and Utilisation*, 2007.

[7] H. Bevrani, F. Habibi, P. Babahajyani, M. Watanabe, and Y. Mitani. Intelligent frequency control in an AC microgrid: Online PSO-based fuzzy tuning approach. *IEEE Transactions on Smart Grid*, 3(4):1935–1944, 2012.

[8] H. Bevrani, T. Ise, and Y. Miura. Virtual synchronous generators: A survey and new perspectives. *International Journal of Electrical Power & Energy Systems*, 54:244–254, 2014.

[9] V. A. Boicea. Energy storage technologies: The past and the present. *Proceedings of the IEEE*, 102(11):1777–1794, 2014.

[10] H. R. Chamorro, I. Riaño, R. Gerndt, I. Zelinka, F. Gonzalez-Longatt, and V. K. Sood. Synthetic inertia control based on fuzzy adaptive differential evolution. *International Journal of Electrical Power & Energy Systems*, 105:803–813, feb 2019.

[11] J. Chen, F. Yang, and Q.-L. Han. Model-free predictive h control for grid-connected solar power generation systems. *IEEE Transactions on Control Systems Technology*, 22(5):2039–2047, 2014.

[12] J. P. Coelho, J. Boaventura-Cunha, and P. B. de Moura Oliveira. *Extended Stability Conditions for CDM Controller Design*. Springer International Publishing, 2015.

[13] S. Curi, D. Gross, and F. Dorfler. Control of low-inertia power grids: A model reduction approach. In *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, 2017.

[14] F. Díaz-González, A. Sumper, O. Gomis-Bellmunt, and R. Villafáfila-Robles. A review of energy storage technologies for wind power applications. *Renewable and Sustainable Energy Reviews*, 16(4):2154–2171, may 2012.

[15] A. Fathi, Q. Shafiee, and H. Bevrani. Robust frequency control of microgrids using an extended virtual synchronous generator. *IEEE Transactions on Power Systems*, 33(6):6289–6297, nov 2018.

[16] A. Fernández-Guillamón, E. Gómez-Lázaro, E. Muljadi, and Á. Molina-García. Power systems with high renewable energy sources: A review of inertia and frequency control strategies over time. *Renewable and Sustainable Energy Reviews*, 115:109369, 2019.

[17] P. F. Frack, P. E. Mercado, and M. G. Molina. Extending the VISMA concept to improve the frequency stability in microgrids. In *2015 18th International Conference on Intelligent System Application to Power Systems (ISAP)*, 2015.

[18] Z.-L. Gaing. A particle swarm optimization approach for optimum design of PID controller in AVR system. *IEEE Transactions on Energy Conversion*, 19(2):384–391, 2004.

[19] M. Garmroodi, G. Verbic, and D. J. Hill. Frequency support from wind turbine generators with a time-variable droop characteristic. *IEEE Transactions on Sustainable Energy*, 9(2):676–684, apr 2018.

[20] I. Goodfellow, Y. Bengio, and A. Courvile. *Deep Learning*. The MIT Press, 2016.

[21] D. Groß and F. Dörfler. On the steady-state behavior of low-inertia power systems. *IFAC-PapersOnLine*, 50(1):10735–10741, 2017.

[22] D. Groß, S. Bolognani, B. K. Poolla, and F. Dörfler. Increasing the resilience of low-inertia power systems by virtual inertia and damping. In *Proceedings of IREP'2017 Symposium*. ETH Zurich, 2017.

[23] S. S. Guggilam, C. Zhao, E. Dall'Anese, Y. C. Chen, and S. V. Dhople. Engineering inertial and primary-frequency response for distributed energy resources. In *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*. IEEE, 2017.

[24] F. Guo, L. Wang, C. Wen, D. Zhang, and Q. Xu. Distributed voltage restoration and current sharing control in islanded DC microgrid systems without continuous communication. *IEEE Transactions on Industrial Electronics*, 67(4):3043–3053, 2020.

[25] W. Guo, F. Liu, J. Si, and S. Mei. Incorporating approximate dynamic programming-based parameter tuning into PD-type virtual inertia control of DFIGs. In *International Joint Conference on Neural Networks*, 2013.

[26] H.-G. Han, L. Zhang, Y. Hou, and J.-F. Qiao. Nonlinear model predictive control based on a self-organizing recurrent neural network. *IEEE Transactions on Neural Networks and Learning Systems*, 27(2):402–415, 2016.

[27] B. Hekimoglu. Optimal tuning of fractional order PID controller for DC motor speed control via chaotic atom search optimization algorithm. *IEEE Access*, 7:38100–38114, 2019.

[28] I. A. Hiskens and E. M. Fleming. Control of inverter-connected sources in autonomous microgrids. In *2008 American Control Conference*. IEEE, 2008.

[29] Y. Hu, W. Wei, Y. Peng, and J. Lei. Fuzzy virtual inertia control for virtual synchronous generator. In *2016 35th Chinese Control Conference (CCC)*, 2016.

[30] S. Jung, Y. T. Yoon, and J.-H. Huh. An efficient micro grid optimization theory. *Mathematics*, 8(4):560, 2020.

[31] J. Kennedy and R. Eberhart. Particle swarm optimization. In *Proceedings of ICNN95 - International Conference on Neural Networks*, volume 4, pages 1942–1948, 1995.

[32] T. Kerdphol, Y. Qudaih, and Y. Mitani. Optimum battery energy storage system using PSO considering dynamic demand response for microgrids. *International Journal of Electrical Power & Energy Systems*, 83:58–66, 2016.

[33] T. Kerdphol, F. Rahman, Y. Mitani, K. Hongesombut, and S. Küfeoğlu. Virtual inertia control-based model predictive control for microgrid frequency stabilization considering high renewable energy integration. *Sustainability*, 9(5):773, 2017.

[34] T. Kerdphol, F. S. Rahman, Y. Mitani, M. Watanabe, and S. Küfeoğlu. Robust virtual inertia control of an islanded microgrid considering high penetration of renewable energy. *IEEE Access*, 6:625–636, 2018.

[35] T. Kerdphol, F. S. Rahman, M. Watanabe, and Y. Mitani. Robust virtual inertia control of a low inertia microgrid considering frequency measurement effects. *IEEE Access*, 7:57550–57560, 2019.

[36] T. Kerdphol, F. S. Rahman, M. Watanabe, and Y. Mitani. *Virtual Inertia Synthesis and Control*. Springer International Publishing, 2021.

[37] T. Kerdphol, F. S. Rahman, M. Watanabe, Y. Mitani, D. Turschner, and H.-P. Beck. Enhanced virtual inertia control based on derivative technique to emulate simultaneous inertia and damping properties for microgrid frequency regulation. *IEEE Access*, 7:14422–14433, 2019.

[38] T. Kerdphol, M. Watanabe, K. Hongesombut, and Y. Mitani. Self-adaptive virtual inertia control-based Fuzzy logic to improve frequency stability of microgrid with high renewable penetration. *IEEE Access*, 7:76071–76083, 2019.

[39] A. A. Khan, M. Q. Khan, S. G. Satti, and M. Adil. Robust control of hybrid distributed generation for frequency regulation. In *2017 14th International Bhurban Conference on Applied Sciences and Technology (IBCAST)*, 2017.

[40] P. Kou, D. Liang, L. Yu, and L. Gao. Nonlinear model predictive control of wind farm for system frequency support. *IEEE Transactions on Power Systems*, 34(5):3547–3561, 2019.

[41] B. Kroposki, B. Johnson, Y. Zhang, V. Gevorgian, P. Denholm, B.-M. Hodge, and B. Hannegan. Achieving a 100% renewable grid: Operating electric power systems with extremely high levels of variable renewable energy. *IEEE Power and Energy Magazine*, 15(2):61–73, 2017.

[42] D. Kumar, B. K. Mukherjee, H. D. Mathur, H. Siguerdidjane, and S. Bhanot. Forecast-based modeling and robust frequency control of standalone microgrids considering high penetration of renewable sources. *International Transactions on Electrical Energy Systems*, 31(2), 2020.

[43] P. Kundur. *Power System Stability and Control*. MCGRAW HILL BOOK CO, 1994.

[44] H.-J. Kunisch, K. G. Kramer, and H. Dominik. Battery energy storage another option for load-frequency-control and instantaneous reserve. *IEEE Transactions on Energy Conversion*, EC-1(3):41–46, 1986.

[45] M. Kusy and R. Zajdel. Application of reinforcement learning algorithms for the adaptive computation of the smoothing parameter for probabilistic neural network. *IEEE Transactions on Neural Networks and Learning Systems*, 26(9):2163–2175, sep 2015.

[46] F. Lamnabhi-Lagarrigue, A. Annaswamy, S. Engell, A. Isaksson, P. Khargonekar, R. M. Murray, H. Nijmeijer, T. Samad, D. Tilbury, and P. V. den Hof. Systems & control for the future of humanity, research agenda: Current and future roles, impact and grand challenges. *Annual Reviews in Control*, 43:1–64, 2017.

[47] Y. Levron and J. Belikov. Control of energy storage devices under uncertainty using nonlinear feedback systems. In *2020 IEEE Power & Energy Society General Meeting (PESGM)*. IEEE, 2020.

[48] 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, 2019.

[49] A. V. Lipatov and N. I. Sokolov. On some sufficient conditions for stability and instability of linear continuous stationary systems. *Avtomat. i Telemekh.*, 39(9):1285–1291, 1978.

[50] L. Liu, F. Pan, and D. Xue. Variable-order fuzzy fractional PID controller. *ISA Transactions*, 55:227–233, 2015.

[51] G. Magdy, H. Ali, and D. Xu. A new synthetic inertia system based on electric vehicles to support the frequency stability of low-inertia modern power grids. *Journal of Cleaner Production*, 297:126595, may 2021.

[52] G. Magdy, A. Bakeer, G. Shabib, A. A. Elbaset, and Y. Mitani. Decentralized model predictive control strategy of a realistic multi power system automatic generation control. In *Nineteenth International Middle East Power Systems Conference*, 2017.

[53] G. Magdy, G. Shabib, A. A. Elbaset, and Y. Mitani. A novel coordination scheme of virtual inertia control and digital protection for microgrid dynamic security considering high renewable energy penetration. *IET Renewable Power Generation*, 13(3):462–474, 2019.

[54] E. Mamdani and N. Baaklini. Prescriptive method for deriving control policy in a fuzzy-logic controller. *Electronics Letters*, 11(25-26):625, 1975.

[55] S. Manabe. Coefficient diagram method as applied to the attitude control of controlled-bias-momentum satellite. *IFAC Proceedings Volumes*, 27(13):327–332, 1994.

[56] S. Manabe. Importance of coefficient diagram in polynomial method. In *42nd IEEE International Conference on Decision and Control (IEEE Cat. No.03CH37475)*, 2003.

[57] S. Manabe. Coefficient diagram method in mimo application: an aerospace case study. *IFAC Proceedings Volumes*, 38(1):7–12, 2005.

[58] S. Mariethoz, A. Fuchs, and M. Morari. A VSC-HVDC decentralized model predictive control scheme for fast power tracking. *IEEE Transactions on Power Delivery*, 29(1):462–471, 2014.

[59] W. Mendieta and C. A. Canizares. Primary frequency control in isolated microgrids using thermostatically controllable loads. *IEEE Transactions on Smart Grid*, 12(1):93–105, 2021.

[60] K. Mentesidi, R. Garde, M. Aguado, and E. Rikos. Implementation of a fuzzy logic controller for virtual inertia emulation. In *2015 International Symposium on Smart Electric Distribution Systems and Technologies (EDST)*, 2015.

[61] F. Milano, F. Dörfler, G. Hug, D. J. Hill, and G. Verbic. Foundations and challenges of low-inertia systems (invited paper). In *2018 Power Systems Computation Conference (PSCC)*, 2018.

[62] M. E. Mokadem, V. Courtecuisse, C. Saudemont, B. Robyns, and J. Deuse. Fuzzy logic supervisor-based primary frequency control experiments of a variable-speed wind generator. *IEEE Transactions on Power Systems*, 24(1):407–417, 2009.

[63] M. G. Molina. *GRID ENERGY STORAGE SYSTEMS*. Wiley, 2019.

[64] H.-J. Moon, J. W. Chang, S.-Y. Lee, and S.-I. Moon. Autonomous active power management in isolated microgrid based on proportional and droop control. *Energy Procedia*, 153:48–55, 2018.

[65] A. Naderipour, Z. Abdul-Malek, M. Hajivand, Z. M. Seifabad, M. A. Farsi, S. A. Nowdeh, and I. F. Davoudkhani. Spotted hyena optimizer algorithm for capacitor allocation in radial distribution system with distributed generation and microgrid operation considering different load types. *Scientific Reports*, 11(1), 2021.

[66] K. Nosrati, V. Skiparev, A. Tepljakov, E. Petlenkov, J. Belikov, and Y. Levron. Constrained intelligent frequency control in an AC microgrid: An online reinforcement learning based PID tuning approach. In *2023 IEEE Power and Energy Society General Meeting*, pages 1–5, 2023.

[67] K. Nosrati, V. Skiparev, A. Tepljakov, E. Petlenkov, Y. Levron, and J. Belikov. Coordinated PI-based frequency deviation control of isolated hybrid microgrid: An online multi-agent tuning approach via reinforcement learning. In *2022 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*, pages 1–5, 2022.

[68] K. Nosrati, V. Skiparev, A. Tepljakov, E. Petlenkov, Y. Levron, and J. Belikov. 9 - application of neural network based variable fractional order pid controllers for load frequency control in isolated microgrids. In D. K. Mishra, L. Li, J. Zhang, and M. J. Hossain, editors, *Power System Frequency Control*, pages 203–216. Academic Press, 2023.

[69] Y. Ojo, J. Watson, and I. Lestas. A review of reduced-order models for microgrids: Simplifications vs accuracy.

[70] P. Ostalczyk and P. Duch. Variable, fractional-order PID controller synthesis novelty method. In W. Wang, editor, *Control Based on PID Framework*, chapter 1. IntechOpen, Rijeka, 2021.

[71] A. Oudalov, D. Chartouni, and C. Ohler. Optimizing a battery energy storage system for primary frequency control. *IEEE Transactions on Power Systems*, 22(3):1259–1266, 2007.

[72] N. U. Padmawansa and L. N. W. Arachchige. Improving transient stability of an islanded microgrid using PV based virtual synchronous machines. In *2020 Moratuwa Engineering Research Conference (MERCon)*. IEEE, 2020.

[73] I. Podlubny. Fractional-order systems and PIλDμ-controllers. *IEEE Transactions on Automatic Control*, 44:208–214, 1999.

[74] Z. Qi, Q. Shi, and H. Zhang. Tuning of digital PID controllers using particle swarm optimization algorithm for a CAN-based DC motor subject to stochastic delays. *IEEE Transactions on Industrial Electronics*, 67(7):5637–5646, 2020.

[75] P. F. Ribeiro and M. L. Crow. Energy storage systems for advanced power applications. *Proceedings of the IEEE*, 89(12):1744–1756, 2001.

[76] D. M. Rosewater, D. A. Copp, T. A. Nguyen, R. H. Byrne, and S. Santoso. Battery energy storage models for optimal control. *IEEE Access*, 7:178357–178391, 2019.

[77] N. Sa-ngawong and I. Ngamroo. Optimal fuzzy logic-based adaptive controller equipped with DFIG wind turbine for frequency control in stand alone power system. In *2013 IEEE Innovative Smart Grid Technologies-Asia (ISGT Asia)*. IEEE, 2013.

[78] J. Schiffer, D. Zonetti, R. Ortega, A. M. Stanković, T. Sezi, and J. Raisch. A survey on modeling of microgrids—from fundamental physics to phasors and voltage sources. *Automatica*, 74:135–150, 2016.

[79] K. Sharifabadi, L. Harnefors, H.-P. Nee, S. Norrga, and R. Teodorescu. *Design, Control, and Application of Modular Multilevel Converters for HVDC Transmission Systems*. John Wiley & Sons, Ltd., first edition, 2016.

[80] D. Shrestha, U. Tamrakar, N. Malla, Z. Ni, and R. Tonkoski. Reduction of energy consumption of virtual synchronous machine using supplementary adaptive dynamic programming. In *IEEE International Conference on Electro Information Technology*, 2016.

[81] D. Silver, G. Lever, N. Heess, T. Degris, D. Wierstra, and M. Riedmiller. Deterministic policy gradient algorithms. In *31st International Conference on Machine Learning*, volume 22, Beijing, China, 2014.

[82] D. Singh and K. Seethalekshmi. A review on various virtual inertia techniques for distributed generation. In *2020 International Conference on Electrical and Electronics Engineering (ICE3)*. IEEE, 2020.

[83] V. P. Singh, S. R. Mohanty, N. Kishor, and P. K. Ray. Robust H-infinity load frequency control in hybrid distributed generation system. *International Journal of Electrical Power & Energy Systems*, 46:294–305, 2013.

[84] V. Skiparev, J. Belikov, and E. Petlenkov. Reinforcement learning based approach for virtual inertia control in microgrids with renewable energy sources. In *IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe)*, The Hague, NL, 2020.

[85] V. Skiparev, J. Belikov, and E. Petlenkov. MIMO reinforcement learning based approach for frequency support in microgrids with high renewable energy penetration. In *IEEE PES General Meeting*, Washington DC, USA, 2021.

[86] V. Skiparev, J. Belikov, E. Petlenkov, and Y. Levron. Reinforcement learning based MIMO controller for virtual inertia control in isolated microgrids. In *IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*, Novi Sad, Serbia, 2022.

[87] V. Skiparev, R. Machlev, N. R. Chowdhury, Y. Levron, E. Petlenkov, , and J. Belikov. Virtual inertia control methods in islanded microgrids. *Energies*, 14(6):1562, 2021.

[88] V. Skiparev, K. Nosrati, J. Belikov, A. Tepljakov, and E. Petlenkov. An enhanced NN-based load frequency control design of MGs: A fractional order modeling method. In *2023 IEEE International Conference on Compatibility, Power Electronics and Power Engineering (CPE-POWERENG)*, Tallinn, Estonia, 2023.

[89] V. Skiparev, K. Nosrati, A. Tepljakov, E. Petlenkov, Y. Levron, J. Belikov, and J. M. Guerrero. Virtual inertia control of an isolated microgrid using NN-VFOPID controller: A self-tuning approach. *IEEE Transactions on Sustainable Energy*, 2023.

[90] N. Sockeel, J. Gafford, B. Papari, and M. Mazzola. Virtual inertia emulator-based model predictive control for grid frequency regulation considering high penetration of inverter-based energy storage system. *IEEE Transactions on Sustainable Energy*, 11(4):2932–2939, 2020.

[91] R. S. Sutton and A. G. Barto. *Reinforcement Learning: An Introduction*. The MIT Press, second edition, 2018.

[92] U. Tamrakar, T. M. Hansen, R. Tonkoski, and D. A. Copp. Model predictive frequency control of low inertia microgrids. In *2019 IEEE 28th International Symposium on Industrial Electronics (ISIE)*, 2019.

[93] U. Tamrakar, D. Shrestha, M. Maharjan, B. Bhattarai, T. Hansen, and R. Tonkoski. Virtual inertia: Current trends and future directions. *Applied Sciences*, 7(7):654, 2017.

[94] A. Tepljakov, E. Petlenkov, and J. Belikov. FOMCON: Fractional order modeling and control toolbox for MATLAB. In *The 18th International Conference Mixed Design of Integrated Circuits and Systems*, pages 684–689, Gliwice, Poland, 2011.

[95] L. Toma, M. Sanduleac, S. A. Baltac, F. Arrigo, A. Mazza, E. Bompard, A. Musa, and A. Monti. On the virtual inertia provision by BESS in low inertia power systems. In *2018 IEEE International Energy Conference (ENERGYCON)*, 2018.

[96] L. Toma, M. Sanduleac, S. A. Baltac, F. Arrigo, A. Mazza, E. Bompard, A. Musa, and A. Monti. On the virtual inertia provision by BESS in low inertia power systems. In *2018 IEEE International Energy Conference (ENERGYCON)*, 2018.

[97] A. Ulbig, T. S. Borsche, and G. Andersson. Impact of low rotational inertia on power system stability and operation. In *The International Federation of Automatic Control IFAC Cape Town*, 2014.

[98] A. Ulbig, T. S. Borsche, and G. Andersson. Impact of low rotational inertia on power system stability and operation. *IFAC Proceedings Volumes*, 47(3):7290–7297, 2014.

[99] T. V. Van, K. Visscher, J. Diaz, V. Karapanos, A. Woyte, M. Albu, J. Bozelie, T. Loix, and D. Federenciuc. Virtual synchronous generator: An element of future grids. In *2010 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT Europe)*, 2010.

[100] B. Vedik, R. Kumar, R. Deshmukh, S. Verma, and C. K. Shiva. Renewable energy-based load frequency stabilization of interconnected power systems using quasi-oppositional dragonfly algorithm. *Journal of Control, Automation and Electrical Systems*, 32(1):227–243, oct 2020.

[101] P. Vorobev, P.-H. Huang, M. A. Hosani, J. L. Kirtley, and K. Turitsyn. A framework for development of universal rules for microgrids stability and control. In *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, 2017.

[102] Y. Wang and G. Zhao. A comparative study of fractional-order models for lithium-ion batteries using runge kutta optimizer and electrochemical impedance spectroscopy. *Control Engineering Practice*, 133:105451, 2023.

[103] Z. Wang, Q. Wang, D. He, Q. Liu, X. Zhu, and J. Guo. An improved particle swarm optimization algorithm based on fuzzy PID control. In *2017 4th International Conference on Information Science and Control Engineering (ICISCE)*, 2017.

[104] L. Yang and Z. Hu. Coordination of generators and energy storage to smooth power fluctuations for multi-area microgrid clusters: A robust decentralized approach. *IEEE Access*, 9:12506–12520, 2021.

[105] K. Y. Yap, C. R. Sarimuthu, and J. M.-Y. Lim. Virtual inertia-based inverters for mitigating frequency instability in grid-connected renewable energy system: A review. *Applied Sciences*, 9(24):5300, 2019.

[106] J. Zeng, B. Zhang, C. Mao, and Y. Wang. Use of battery energy storage system to improve the power quality and stabilty of wind farms. In *2006 International Conference on Power System Technology*, 2006.

[107] Q.-C. Zhong, G. C. Konstantopoulos, B. Ren, and M. Krstic. Improved synchronverters with bounded frequency and voltage for smart grid integration. *IEEE Transactions on Smart Grid*, 9(2):786–796, 2018.

[108] Q.-C. Zhong and G. Weiss. Synchronverters: Inverters that mimic synchronous generators. *IEEE Transactions on Industrial Electronics*, 58(4):1259–1267, 2011.

[109] J. Zhu, Y. Zheng, Y. Wang, and Y. Yuan. Isolated microgrid capacity configuration considering economic risk of customer interruption. In *2019 IEEE Innovative Smart Grid Technologies - Asia (ISGT Asia)*. IEEE, 2019.

[110] M. F. Zia, E. Elbouchikhi, and M. Benbouzid. Microgrids energy management systems: A critical review on methods, solutions, and prospects. *Applied Energy*, 222:1033–1055, 2018.

Acknowledgements

Abstract
Virtual inertia control of microgrids using deep reinforcement learning methods

Concerns about climate change turned the development of power systems towards the direction of increased energy generation by renewable energy sources and decentralisation of the grid. However, increased penetration of renewable energy sources triggered the phenomenon known as low inertia. A power system with decreased inertia creates many problems and risks in relation to total power stability, including decreased frequency stability, frequency mismatch between parts of the interconnected grid, power consumption/production imbalance and power outrage. In order to increase the inertia and mitigate risks, many engineers propose to emulate the inertia of the synchronous machine by power electronics-based devices known as virtual synchronous generators (VSG). However, emulation of virtual inertia can only be accurate if its operation is controlled by a robust closed-loop algorithm. In the last few years, many classical, hybrid and advanced algorithms have tried to emulate optimal inertia; however, most of them still have uncovered limitations in scalability and deployment for decentralised power systems, limited flexibility and computational complexity. In other words, they cannot mitigate all potential stability risks in the full-scale decentralised and interconnected grid. In this doctoral dissertation, we propose a deep reinforcement learning (RL)-based method with a new reward/punishment system. The result of the proposed advanced controller is increased efficiency of inertia emulation in scenarios with different system inertia, renewable energy and domestic loads (dis)connection. In addition, we demonstrate scalability, since RL can train an artificial neural network (ANN) as an independent controller and as a hybrid combined with V(FO)PID controllers and a control scheme of advanced energy storage. Moreover, we demonstrate deployment potential in microgrid applications, e.g. provide coordinated control in isolated hybrid microgrid using multi-agent architecture.

Kokkuvõte
Mikrovõrkude virtuaalse inertsi juhtimine sügava stiimulõppe meetoditega

Kliimamuutustega seotud probleemid on suunanud elektrisüsteemide arengu taastuvate energiaallikate suurema energiatootmise ja võrgu detsentraliseerimise suunas. Taastuvate energiaallikate kasvav levik käivitas aga nähtuse, mida tuntakse kui madalat inertsi. Vähenenud inertsiga elektrisüsteem tekitab mitmeid probleeme ja riske seoses võrgu stabiilsusega, sealhulgas vähenenud sagedusstabiilsus, sageduse mittevastavus võrku ühendatud osade vahel, elektritarbimise ja -tootmise tasakaalustamatus ning elektrikatkestused. Inertsi suurendamiseks ja riskide vähendamiseks teevad paljud insenerid ettepaneku emuleerida sünkroonmasina inertsi jõuelektroonikapõhiste seadmetega, mida nimetatakse virtuaalseteks sünkroongeneraatoriteks (VSG). Virtuaalse inertsi emuleerimine saab aga olla täpne ainult siis, kui selle tööd juhib robustne suletud ahela algoritm. Viimastel aastatel on paljud klassikalised, hübriid- ja täiustatud algoritmid püüdnud emuleerida optimaalset inertsi; enamikul neist on siiski leitud piirangud seoses skaleeritavuse ja detsentraliseeritud elektrisüsteemide kasutuselevõtuga, piiratud paindlikkusega ja arvutuskompleksusega. Teisisõnu, nad ei suuda leevendada kõiki võimalikke stabiilsusriske täiemahulises detsentraliseeritud ja ühendatud võrgus. Käesolevas doktoritöös pakume välja sügava stiimulõppel (RL) põhineva meetodi koos uue tasu/karistuse süsteemiga. Väljapakutud täiustatud kontrolleri tulemuseks on suurem tõhusus inertsi jäljendamisel stsenaariumides, kus on erinev süsteemi inerts, taastuvenergia ja kodumaiste koormuste (lahti)ühendamine. Lisaks demonstreerime skaleeritavust, kuna RL saab õpetada tehisnärvivõrgu (ANN) iseseisva kontrollerina ja hübriidina koos V(FO)PID-regulaatorite ja täiustatud energiasalvestuse juhtimisskeemiga. Lisaks sellele demonstreerime kasutuselevõtu potentsiaali mikrovõrkude rakendustes, nt koordineeritud juhtimise tagamine isoleeritud hübriidmikrovõrgus, kasutades mitmeagendi arhitektuuri.

Appendix 1

IX

V. Skiparev, J. Belikov, and E. Petlenkov. Reinforcement learning based approach for virtual inertia control in microgrids with renewable energy sources. In *IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe)*, The Hague, NL, 2020

Reinforcement Learning based Approach for Virtual Inertia Control in Microgrids with Renewable Energy Sources

Vjatseslav Skiparev, Juri Belikov
Department of Software Science
Tallinn University of Technology
Akadeemia tee 15a, 12618 Tallinn
Email: {vjatseslav.skiparev, juri.belikov}@taltech.ee

Eduard Petlenkov
Department of Computer Systems
Tallinn University of Technology
Akadeemia tee 15a, 12618 Tallinn
Email: eduard.petlenkov@taltech.ee

Abstract—**The increasing penetration of distributed and renewable energy sources into power grids results in various control challenges. Among others is a problem of decreasing inertia, especially in the case of islanded microgrids with high penetration of a low-inertia power sources. One approach is based on a concept of additional virtual inertia control. In this paper we propose reinforcement learning based virtual inertia control with deep deterministic policy gradients based optimization algorithm. The proposed solution is demonstrated using standard topology of a microgrid and compares to H-infinity and optimally tuned PI controllers.**

Index Terms—**virtual inertia, reinforcement learning, microgrids, renewable energy**

I. Introduction

The last decade has witnessed the raising concerns regarding continuous environmental changes. This has motivated various actions for shifting energy production toward more intense utilization of renewable energy sources such as wind or solar. However, due to the natural lack of inertia, massive switch to renewable energy creates multiple stability related problems in the grid. These challenges resulted in various solutions; among others different control algorithms for the optimal virtual inertia control [1].

Several recent works [2], [3] have addressed the problem of optimal frequency regulation with high penetration of renewable energy sources. In [4] the robust H_∞ controller was developed for the Rate of Change of Frequency (RoCoF) stability support. The proposed solution has shown advantages over conventional virtual inertia control and optimally tuned PI controller in scenarios when the wind farm is connected, solar panels are disconnected, and the system inertia is 100% and 10%. In [5] this problem was further studied and virtual inertia control was implemented using Fuzzy-logic based approach. The proposed algorithm performed robustly under different scenarios with additional uncertainties, including 80%, 40%, 30% system inertia and mismatch in primary/secondary control loops. In [6] the model predictive control scheme was proposed and compared to Fuzzy-logic controller during sudden load changes. Unlike the previous works, the studied microgrid has conceptual differences such as closed-loop turbine system,

renewable energy power generation from two more complex wind farm models, and minor differences in turbine and system inertia transfer functions. Similar ideas are presented in [7], but without renewable energy penetration. In [8] optimization of PI controller was performed by Particle Swarm Optimization technique combined with digital frequency protection system in scenarios of (dis)connecting load and renewable energy sources.

In this paper we join the current trend of research and propose an effective deep Reinforcement learning (RL) algorithm for the robust virtual inertia control in microgrids with varying system inertia. The approach is designed to control the frequency using Deep Deterministic Policy Gradient (DDPG) in Neural Actor-Critic strategy. Specifically, we design reinforcement learning control to adjust the frequency deviation using typical microgrid model with several fluctuating renewable energy sources. The reward system has been adjusted by changing the error band for controller reward/punishment, selecting the optimal constants at each step, and modifying the optimal number of fully connected layers in the network. The proposed solution is compare to H_∞ and optimally tuned PI controllers using several statistical measures. It was observed that on average the proposed solution performs better.

II. Motivation and Problem Statement

The frequency stability problem forms a core challenge for the future power systems. In the light of increasing penetration of renewable energy sources and decreasing inertia, the development of new robust methodologies with potential advantages for handling low-inertia uncertainties and domestic load non-linearities forms considerable challenge. Implementations of traditional control methods have certain limitations for optimal management of virtual inertia. For example, Fuzzy-logic based controller performance depends on the designed landscape pattern emerging from fuzzy rules. H_∞ controller is usually based on advance optimization procedure. In contrast, reinforcement learning based algorithms do not demand construction of the specific representation of a process and have potential for online optimization. The low-

inertia grid stability support based on reinforcement learning with neural actor-critic architecture was addressed in several recent works. For example, in [9] approach was combined with approximate dynamic programming and applied to tune PD controller for virtual inertia control in power system with parallel connected three synchronous generators and one wind farm. In [10] reinforcement learning was combined with dynamic programming and used to tune PID controller for virtual synchronous machine. The work [11] proposed neural actor-critic with adaptive learning approach for voltage/current regulation in autonomous microgrids. In this paper we utilize deep deterministic policy gradient method introduced in [12]. Compared to other reinforcement learning optimization methods oriented to high-dimensional and discrete time spaces, DDPG was developed for low-dimensional and continuous space actions, which appears to be suitable for the addressed scenario. This optimization algorithm was previously tested for pendulum and cart-pole balance, cheetah acceleration and moving gripper position control [12].
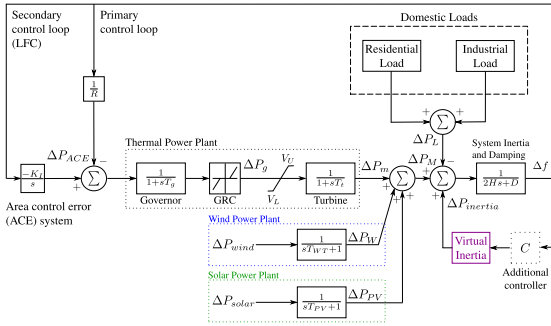
Fig. 1. Dynamic model of studied microgrid with hierarchical control.

A. Structure of the Studied Microgrid

The studied microgrid is adopted from several recent publications [2], [4], [5], [8] and depicted in Fig. 1. The addressed scenario includes residential/industrial loads, energy sources (thermal power plant, wind farm, and solar power plant), and energy storage system. Thermal power plant is composed of a governor with generator rate constraint (GRC) and turbine with frequency rate limiter, which restricts the valve opening/closing (V_U, V_L). The dynamic model of microgrid utilizes the hierarchical architecture with primary and secondary control loops. The primary control loop has droop coefficient $1/R$, and the secondary loop has an area control error system (ACE) with the second frequency controller K_I and the first-order integrator. The frequency regulation is performed by virtual inertia device with additional optional controller. Microgrid balancing system performed as transfer function with microgrid damping coefficient D and system inertia H. The power generation by variable energy sources computed as combination of two random signals with first-order holder.

TABLE I
NOMENCLATURE: PARAMETERS OF MICROGRID

Variable	Physical meaning
ΔP_m	generated power change from the distributed generator
T_t	time constant of the turbine
ΔP_g	governor valve-position change
T_g	time constant of the governor
ΔP_{ACE}	control signal change for secondary control
K_I	integral control variable gain
ΔP_W	change in generated power-based wind farm
ΔP_{wind}	initial wind power variation
T_{WT}	time constant of wind turbines
ΔP_{PV}	change in generated power-based solar farm
ΔP_{solar}	initial solar power variation
T_{PV}	time constant of the solar system
ΔP_L	load power change
ΔP_{RL}	variations in residential loads
ΔP_{IL}	variations in industrial loads

The deviation of frequency in the studied microgrid can be calculated as

$$\Delta f = \frac{1}{2Hs + D}(\Delta P_m + \Delta P_W + \Delta P_{PV} + \Delta P_{inertia} - \Delta P_L), \quad (1)$$

where

$$\Delta P_m = \frac{1}{1 + sT_t}\Delta P_g,$$

$$\Delta P_g = \frac{1}{1 + sT_g}\left(\Delta P_{ACE} - \frac{1}{R}\Delta f\right),$$

$$\Delta P_{ACE} = \frac{K_I}{s}\Delta f,$$

$$\Delta P_W = \frac{1}{1 + sT_{WT}}\Delta P_{wind},$$

$$\Delta P_{PV} = \frac{1}{1 + sT_{PV}}\Delta P_{solar},$$

$$\Delta P_L = \Delta P_{RL} + \Delta P_{IL},$$

and parameters are summarized in Table I.

B. Modeling of Virtual Inertia System

Virtual synchronous generator (VSG) is considered as a converting power alternative to a real synchronous machine [13]. Such generator can be applied in systems with high level of fluctuating renewable power to enhance the frequency stability in conditions of dynamic power generation. Virtual inertia (VI) is a specific part of VSG designed to compensate the lack of inertia [1]. Unfortunately, the default operational limitations of virtual inertia device cannot provide enough frequency support. Therefore, the additional robust controller has to be used to deal with nonlinearities in low-inertia environments. Traditionally, virtual inertia consists of derivative component, designed controller, energy storage system and power limiter as depicted in Fig. 2. The inertia power saturation limiter ($P_{i,\max}$, $P_{i,\min}$) provides the additional robustness for tested algorithms and creates limitations for more realistic simulations.

Fig. 2. Typical structure of virtual inertia block.

The main concept used in virtual inertia control is the so-called Rate of Change of Frequency (RoCoF), which can be calculated as

$$\text{RoCoF} = \frac{\mathrm{d}(\Delta f)}{\mathrm{d}t}. \tag{2}$$

The RoCoF defines time derivative of the frequency signal, which is used to calculate the inertia response of a system as

$$\Delta P_{inertia} = \frac{K_{VI}}{1 + sT_{VI}} \frac{\mathrm{d}(\Delta f)}{\mathrm{d}t}, \tag{3}$$
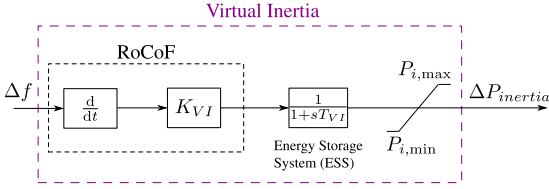
where K_{VI} is the virtual inertia constant, and T_{VI} is time constant of the virtual inertia transfer function.

III. PROPOSED SOLUTION

A. Neural Actor-Critic Algorithm

Neural actor-critic implies the combination of two Artificial Neural Networks: The actor network $\mu(s \mid \theta_\mu)$ and the critic network $Q(s, a \mid \theta_Q)$. In the proposed strategy critic network is supervising algorithm, which tracks errors from interaction of actor a_t with environment s_t according to defined policy. The network corrects them in order to find an optimal estimation of actor action a'_t, which predict max possible reward r, see [14]. The key advantage of reinforcement learning algorithms is the study based on interaction with environment [14]. In other words, when agent makes the action, it expects to get the reward $+r$ or the punishment $-r$. The mechanism of control can be briefly summarized as follows. The measured frequency difference Δf produces the control error which goes as an observation to RL agent. At the same time calculated error goes to the block "calculate reward" to reward or punish neural RL actor. General structure of the proposed controller is presented in Fig. 3.
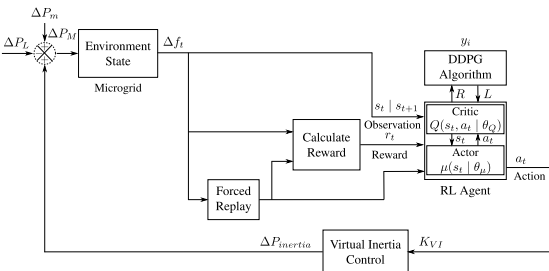
Fig. 3. Structure of implemented reinforcement learning based controller.

B. Deep Deterministic Policy Gradient

Here, we briefly explain the key application details of the proposed controller based on reinforcement learning algorithm optimized by deep deterministic policy gradients. DDPG is model-free reinforcement off-policy learning algorithm, designed for tasks with low-dimensional continuous action space [12]. DDPG study introduces the fusion of deep Q-learning (DQL) and deterministic policy gradient (DPG), which inherits from DQL strategy neural actor-critic. The principle of optimization is based on the search of a minimal difference between target action-value function of the policy y_i and the critic network $Q(s_i, a_i \mid \theta_Q)$ at each actor network $\mu(s \mid \theta_\mu)$ decision a_t (K_{VI}) per step i in the state s_t (i.e., Δf_t). In order to minimize the loss function L and receive the max possible reward r_t per training episode. Next, we summarize the overall procedure in the form of a pseudo-code as shown in Algorithm 1 adapted from [12].

Algorithm 1 DDPG Algorithm

1: Initialize critic $Q(s, a \mid \theta_Q)$ and actor $\mu(s \mid \theta_\mu)$ networks with random weights θ_Q and θ_μ.
2: Initialize target network Q' and μ' with $\theta_{Q'} \leftarrow \theta_Q$, $\theta_{\mu'} \leftarrow \theta_\mu$.
3: Initialize replay buffer R.
4: **for** $episode = 1$ to M **do**
5:     Receive initial process observation as state s_1.
6:     **for** $t = 1$ to T **do**
7:         Select action $a_t = \mu(s_t \mid \theta_\mu)$ according to current policy and disturbances exploration.
8:         Execute action a_t. Observe reward r_t, state s_{t+1}.
9:         Store transition (s_t, a_t, r_t, s_{t+1}) in R.
10:        Sample random minibatch of N transitions (s_i, a_i, r_i, s_{i+1}) from R.
11:        Set $y_i = r_i + \gamma Q'(s_{i+1}, \mu'(s_{i+1}|\theta_{\mu'}) \mid \theta_{Q'})$.
12:        Update critic by minimizing loss: $L = \frac{1}{N} \sum_i (y_i - Q(s_i, a_i \mid \theta_Q))^2$.
13:        Update actor policy using sampled gradient:

$$\nabla_{\theta_\mu} J \approx \frac{1}{N} \sum_i \nabla_a Q\left(s, a \mid \theta^Q\right)|_{s=s_i, \alpha=\mu(s_i)}$$
$$\times \nabla_{\theta_\mu} \mu(s \mid \theta^\mu)|_{s_i}.$$

14:        Update the target network: $\theta^{Q'} \leftarrow \tau\theta^Q + (1-\tau)\theta^{Q'}$ and $\theta^{\mu'} \leftarrow \tau\theta^\mu + (1-\tau)\theta^{\mu'}$.
15:     **end for**
16: **end for**

IV. SIMULATION RESULTS

Consider the microgrid depicted in Fig. 1, which is implemented in MATLAB/Simulink environment. We compare the proposed RL-based control algorithm with the traditional PI controller (with $K_p = 0.25$, $K_i = 0.01$) and H_∞ control. Simulations intend to illustrate the reaction of each controller on RoCoF rise after synchronous connection of thermal power station, wind and solar power plants. Parameters used in

simulations are defined as: $R = 2.4$, $T_g = 0.1$, $T_t = 0.4$, $K_I = 0.05$, GRC $\in [-0.12, 0.12]$, $V_L = -0.5$, $V_U = 0.5$, $T_{WT} = 1.85$, $T_{PV} = 1.5$, $P_{i,\max} = 0.25$, $P_{i,\min} = -0.25$. Moreover, wind power generation varies as $\Delta P_{wind} \in [0.03, 0.05]$, solar power varies as $\Delta P_{solar} \in [0.05, 0.07]$, residential and industrial load profiles are defined by $\Delta P_{RL} \in [0.125, 0.25]$ and $\Delta P_{IL} \in [0.15, 0.25]$ as depicted in Fig. 4. Finally, the stored energy changes in the range $[-0.25, 0.25]$ and energy generated by the thermal power is defined by $\Delta P_g \in [-0.5, 0.5]$. All signals are specified in per-unit system. We consider three different scenarios with high (100%, $H = 0.083$), decreased (80%, $H = 0.067$), and low (40%, $H = 0.032$) inertia. Results of simulations for the nominal case and the the entire period of $T = 1000$ s are shown in Fig. 5. Other cases for a shorter period of $[20, 60]$ are presented in Fig. 6. Performance of control algorithms is validated in steady-state by root mean square error RMSE $= \sqrt{1/n \sum_{i=1}^{n} \Delta f_i^2}$, mean absolute error MAE $= 1/n \sum_{i=1}^{n} |\Delta f_i|$, max/min deviation of frequency signal $\max(\Delta f)$, $\min(\Delta f)$, and integral absolute error IAE $= \int_0^T |\Delta f| \mathrm{d}t$ as summarized in Tables II-V. We remove the first 20 seconds from simulation results, which correspond to the transient process.

Nominal case: The top plot in Fig. 6 shows that the case without additional virtual inertia control and all control algorithms provide almost identical results. This is further emphasized by statistical measures shown in Tables II-V. The proposed algorithm and H_∞ based approach both appear to be the most accurate. However, their advantage is marginal with respect to PI and default controller cases. Therefore, in systems with high inertia utilization of additional virtual inertia controllers is rather optional.

Scenario I: The middle plot in Fig. 6 shows very similar picture to the one obtained in the nominal case. The statistical measures confirm this observation, and the only difference appears for MAE in Table III, where the proposed controller and no virtual inertia control cases become slightly more accurate than H_∞ controller.

Scenario II: The bottom plot in Fig. 6 corresponds to the low-inertia case, and shows the real advantages of additional virtual inertia controllers, since the nominal controller cannot provide satisfactory behavior and becomes oscillating. The statistical measures again show that in most cases the proposed controller becomes slightly more accurate than other techniques. The only case in which H_∞ controller becomes more advantageous is maximum and minimum frequency deviations as indicated in Table IV.

To conclude, the proposed reinforcement learning based additional virtual inertia controller appears to be more accurate in most cases. However, the difference in most cases is very small. At the same time, observe that on average the proposed controller performs better than other controllers in all scenarios as defined by mean average error. This is further confirmed by another related integral absolute error measure.

Table VI provides the performance summary for algorithms with variable and fixed step of the solver. We run the simulation for 10 times and calculated the mean value. It can
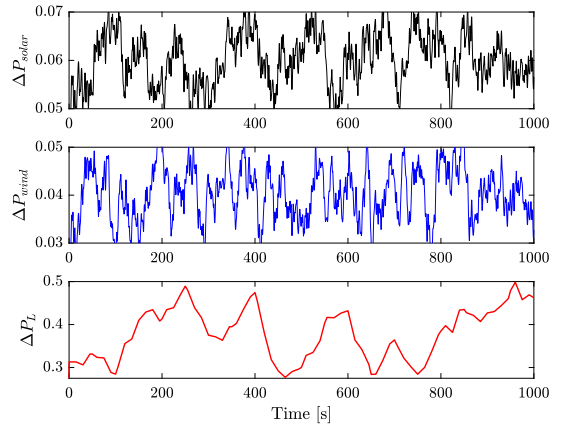
Fig. 4. Power generated by wind and solar plants (top and middle). Domestic load profile (bottom).
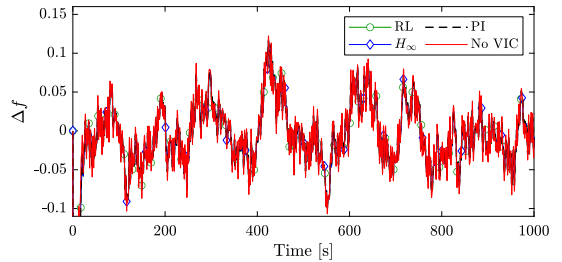
Fig. 5. Frequency variation for the nominal case.
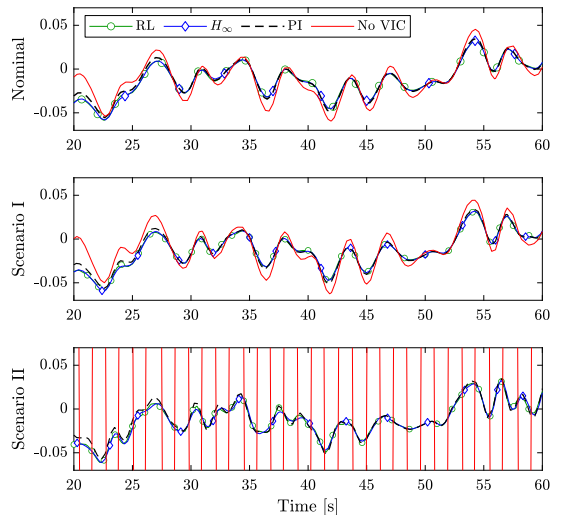
Fig. 6. Frequency variation for inertia coefficients $H = 0.083$ (upper, nominal), $H = 0.067$ (middle, Scenario I), and $H = 0.032$ (bottom, Scenario II). Signals correspond to the proposed algorithm ('—○—'), H_∞ algorithm ('—◇—'), PI controller ('- -'), and the case without additional virtual inertial controller ('—').

TABLE II
PERFORMANCE IN DIFFERENT SCENARIOS: RMSE

Controller	Nominal	Scenario I	Scenario II
Proposed	0.0358	0.0357	**0.0353**
H_∞	**0.0356**	**0.0356**	0.0355
PI	0.0364	0.0366	0.0362
No VIC	0.0379	0.0359	0.6727

TABLE III
PERFORMANCE IN DIFFERENT SCENARIOS: MAE

Controller	Nominal	Scenario I	Scenario II
Proposed	**0.0287**	**0.0285**	**0.0282**
H_∞	**0.0287**	0.0288	0.0288
PI	0.0293	0.0303	0.0292
No VIC	0.0303	**0.0285**	0.5953

TABLE IV
PERFORMANCE IN DIFFERENT SCENARIOS: MAXIMUM AND MINIMUM
FREQUENCY DEVIATION $(\max(\Delta f); \min(\Delta f))$

Controller	Nominal	Scenario I	Scenario II
Proposed	**0.1142**; **−0.0988**	**0.1139**; **−0.0987**	0.1142; −0.0965
H_∞	0.1147; −0.0989	0.1145; −0.0993	**0.1122**; **−0.0960**
PI	0.1162; −0.1011	0.1164; −0.1016	0.1134; −0.0977
No VIC	0.1223; −0.1069	0.1239; −0.1103	1.1733; −1.2097

TABLE V
PERFORMANCE IN DIFFERENT SCENARIOS: INTEGRAL ABSOLUTE ERROR

Controller	Nominal	Scenario I	Scenario II
Proposed	**28.4136**	**28.4259**	**28.3313**
H_∞	28.4461	28.4520	28.4087
PI	28.7557	28.7941	28.7693
No VIC	29.6323	36.6168	594.3056

be seen while in case of variable step H_∞ yields the fastest simulation time, it fails to converge for the fixed step. PI controller results in slightly slower performance. The case without VIC appears to be the fastest for fixed step. Finally, the proposed algorithm works slower than others, due to complex and non-linear structure defined by the Deep Neural Network.

TABLE VI
COMPARISON OF ALGORITHMS FOR VARIABLE AND FIXED STEP

	Elapsed time	
Controller	Variable	Fixed
Proposed	2975.5	26814
H_∞	**613.26**	—
PI	820.88	15227
No VIC	1175.5	**13380**

V. CONCLUSION AND DISCUSSION

Increasing penetration of renewable sources based power creates multiple serious challenges related to the frequency stability of future microgrids. These problems can be solved by adding optimally designed control algorithms. The recent studies indicate that without extra virtual inertia support the system is stable only for the relatively high system inertia. Otherwise, the virtual inertia device is not able to provide satisfactory frequency stability with high penetration of renewable energy sources. Therefore, in this paper we propose the reinforcement learning based control of virtual inertia, and compare to H_∞ and PI controllers. In most situations the proposed controller has shown slightly better results. More complex scenarios and detailed analysis will form subject for the future research. The reinforcement learning is recently introduced optimization strategy. We plan to extend the DDPG Reinforcement learning based controller via modification of the reward/punishment system and derive the controller for MIMO case. Finally, we will apply the proposed technique to study problems of energy transmission and demand side management fields.

VI. ACKNOWLEDGMENT

REFERENCES

[1] F. Milano, F. Dörfler, G. Hug, D. J. Hill, and G. Verbic, "Foundations and challenges of low-inertia systems (invited paper)," in *Power Systems Computation Conference*, 2018.

[2] H. Ali, G. Magdy, B. Li, G. Shabib, A. A. Elbaset, D. Xu, and Y. Mitani, "A new frequency control strategy in an islanded microgrid using virtual inertia control-based coefficient diagram method," *IEEE Access*, vol. 7, pp. 16 979–16 990, 2019.

[3] G. Magdy, A. Bakeer, G. Shabib, A. A. Elbaset, and Y. Mitani, "Decentralized model predictive control strategy of a realistic multi power system automatic generation control," in *Nineteenth International Middle East Power Systems Conference*, 2017.

[4] T. Kerdphol, F. S. Rahman, M. Watanabe, and Y. Mitani, "Robust virtual inertia control of a low inertia microgrid considering frequency measurement effects," *IEEE Access*, vol. 7, pp. 57 550–57 560, 2019.

[5] T. Kerdphol, M. Watanabe, K. Hongesombut, and Y. Mitani, "Self-adaptive virtual inertia control-based fuzzy logic to improve frequency stability of microgrid with high renewable penetration," *IEEE Access*, vol. 7, pp. 76 071–76 083, 2019.

[6] T. Kerdphol, F. Rahman, Y. Mitani, K. Hongesombut, and S. Küfeoğlu, "Virtual inertia control-based model predictive control for microgrid frequency stabilization considering high renewable energy integration," *Sustainability*, vol. 9, no. 5, p. 773, 2017.

[7] U. Tamrakar, T. M. Hansen, R. Tonkoski, and D. A. Copp, "Model predictive frequency control of low inertia microgrids," in *International Symposium on Industrial Electronics*, 2019.

[8] G. Magdy, G. Shabib, A. A. Elbaset, and Y. Mitani, "A novel coordination scheme of virtual inertia control and digital protection for microgrid dynamic security considering high renewable energy penetration," *IET Renew. Power Gen.*, vol. 13, no. 3, pp. 462–474, 2019.

[9] W. Guo, F. Liu, J. Si, and S. Mei, "Incorporating approximate dynamic programming-based parameter tuning into PD-type virtual inertia control of DFIGs," in *International Joint Conference on Neural Networks*, 2013.

[10] D. Shrestha, U. Tamrakar, N. Malla, Z. Ni, and R. Tonkoski, "Reduction of energy consumption of virtual synchronous machine using supplementary adaptive dynamic programming," in *International Conference on Electro Information Technology*, 2016.

[11] M. S. Mahmoud, M. Abouheaf, and A. Sharaf, "Reinforcement learning control approach for autonomous microgrids," *Int. J. Model. Simul*, pp. 1–10, aug 2019.

[12] 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*, 2015. [Online]. Available: https://arxiv.org/abs/1509.02971

[13] H.-P. Beck and R. Hesse, "Virtual synchronous machine," in *International Conference on Electrical Power Quality and Utilisation*, 2007.

[14] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, 2nd ed. The MIT Press, 2018.

Appendix 2

VIII

V. Skiparev, J. Belikov, and E. Petlenkov. MIMO reinforcement learning based approach for frequency support in microgrids with high renewable energy penetration. In *IEEE PES General Meeting*, Washington DC, USA, 2021

MIMO Reinforcement Learning based Approach for Frequency Support in Microgrids with High Renewable Energy Penetration

Vjatseslav Skiparev, Juri Belikov
Department of Software Science
Tallinn University of Technology
Akadeemia tee 15a, 12618 Tallinn, Estonia
{vjatseslav.skiparev, juri.belikov}@taltech.ee

Eduard Petlenkov
Department of Computer Systems
Tallinn University of Technology
Akadeemia tee 15a, 12618 Tallinn, Estonia
eduard.petlenkov@taltech.ee

Abstract—**In this paper we propose here a nonlinear control scheme for frequency support in low-inertia microgrids with high level integration of renewable energy sources. We first develop a multi-loop reinforcement learning based controller with deep deterministic policy gradient optimization. Then, we apply it to the simultaneous frequency support and control of renewable energy generation. In addition, we adjust the reward system to track the thermal power and provide the balance between energy generation and consumption. This modified controller is shown to work well in several practical scenarios, in which it is compared to a single loop RL controller.**

Index Terms—**MicroGrid, Reinforcement learning, Renewable energy**

I. INTRODUCTION

Concerns on future state of environment gained attention to electronic-based renewable energy sources. However, the integration of renewable energy in large amounts creates power stability related problems such as frequency drop, power oscillations and mismatch in energy generation [1]–[3]. These challenges raise important problems in the light of future power systems. Among others frequency support gains an increased trend [3]–[6].

Several recent works [7]–[12] addressed the problem of robust frequency regulation with high penetration of renewable energy sources. For example, in [9] the robust H_∞ controller combined with phase-lock loop (PLL) for microgrid support with 100% and 10% inertia, and 80% penetration renewable energy. In [13] the problem was studied with using fuzzy-logic controller in scenarios with various uncertainties, including 20% and 80% integration, and 80%, 40%, 30% system inertia mismatch in primary/secondary control loops. In [14] the model predictive control approach proposed and compared with fuzzy-logic controller with 34% integration of RES power. In [12] reinforcement learning based control strategy proposed compared with H_∞ and PID in microgrid with 20% of renewable energy integration and 100%, 80%, 40% system inertia. Also frequency control by reinforcement learning strategy was addressed in [15], [16].

Optimal frequency support in islanded microgrids with high integration level of renewable energy sources and decreasing inertia is one of the core challenges for future power systems [17], [18]. Therefore, development of novel and advanced methodologies with potential to handle renewable power, domestic loads and low-inertia uncertainties becomes extremely important. Implementations of traditional control schemes have certain limitations for the optimal frequency support. For example, the standard PID controller is single input single output system with limited performance, including weak tolerance to disturbances and efficiency in specific conditions. Another popular H_∞ controller operates with additional inputs as disturbances and synthesized controller provides input/output relation similar to that of PID [9]. At the same time, data-driven algorithms can be implemented as multi input systems. For example, in [13] fuzzy-logic controller designed with two inputs Δf and ΔP_{RES}. Similar approach is applied in [14] to develop state-space based model predictive control with inputs Δf and ΔP_W. The work [12] proposed reinforcement learning based control approach, which consider the Δf dynamics and acceleration, and total power dynamics ΔP_M. The above mentioned work consider multi-input single-output scenario.

In this paper, we extend our previous work [12] and propose reinforcement learning controller with multi-input multi-output architecture. The new design enables more robust multi-loop frequency support in islanded/isolated microgrids with varying system inertia. Specifically, we develop a technique which provides simultaneous control of renewable energy flow and virtual inertia emulation. Furthermore, we adjust the reward system by changing the error band for controller reward/punishment, selecting the optimal constants at each step, and change the optimal number of artificial neurons and fully connected layers in the network. The benefit of such modification is the possibility to decrease the negative influence of renewable energy and provide the balance in energy production between different energy generation sources. As result extended RL controller provides the effective frequency stability with integration up to 50% of renewable power.

II. STRUCTURE OF THE STUDIED MICROGRID

The studied microgrid is adopted from several recent publications [8]–[10], [13] and depicted in Fig. 1.
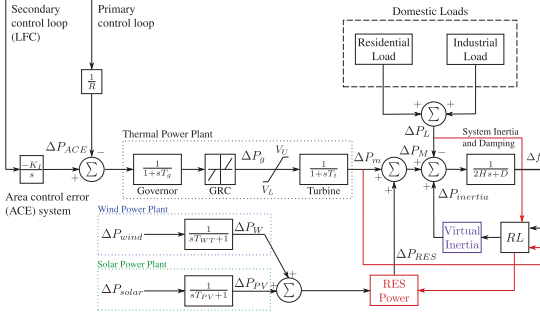

Fig. 1. Block-scheme of the islaned microgrid with hierarchical control.

The addressed setup includes residential ΔP_{RL} and industrial loads ΔP_{IL}, energy sources (thermal power plant ΔP_m, wind farm ΔP_W, and solar power plant ΔP_{PV}), and energy storage system. Thermal power plant is composed of a governor with generator rate constraints ($P_{g,\max}$, $P_{g,\min}$) and turbine with frequency rate limiter, which restricts the valve opening/closing (V_U, V_L). The dynamic model of microgrid utilizes the hierarchical architecture with primary and secondary control loops. The primary control loop has droop coefficient $1/R$, and the secondary loop has an area control error system ΔP_{ACE} with the second frequency controller K_I and the first-order integrator $1/s$. The frequency regulation is performed by virtual inertia $\Delta P_{inertia}$ and renewable energy control ΔP_{RES}. Microgrid balancing system performed as transfer function with microgrid damping coefficient D and system inertia H. The power generation by variable energy sources computed as combination of two random signals with first-order holder. The deviation of frequency in the studied microgrid can be calculated as

$$\Delta f = \frac{1}{2Hs + D}(\Delta P_m + \Delta P_{RES} + \Delta P_{inertia} - \Delta P_L),$$

(1)

where

$$\Delta P_m = \frac{1}{1 + sT_t}\Delta P_g,$$
$$\Delta P_g = \frac{1}{1 + sT_g}\left(\Delta P_{ACE} - \frac{1}{R}\Delta f\right),$$
$$\Delta P_{ACE} = \frac{K_I}{s}\Delta f,$$
$$\Delta P_L = \Delta P_{RL} + \Delta P_{IL}.$$

Modeling parameters of the microgrid are summarized in Table I.

TABLE I
NOMENCLATURE: PARAMETERS OF MICROGRID.

Parameter	Value
T_t(s), time constant of the turbine	0.4
T_g(s), time constant of the governor	0.1
K_I(s), integral control variable gain	0.05
H(p.u.MW s), System inertia	0.083
D(p.u.MW s), System damping coefficient	0.05
R(Hz/p.u.MW), Droop characteristic	2.4
T_{WT}(s), time constant of wind turbines	1.85
T_{PV}(s), time constant of the solar system	1.5
T_{RES}(s), time constant of the renewable energy storage	1
T_{VI}(s), time constant of the virtual inertia emulation	10
V_U, Maximum limit of valve gate speed	0.5
V_L, Minimum limit of valve gate speed	-0.5
$P_{inertia,\max}$, Maximum capacity of ESS	0.25
P_{inetia_min}, Minimum capacity of ESS	-0.25
$P_{RES,\max}$, Maximum capacity of renewable power storage	0.5
$P_{RES,\min}$, Minimum capacity of renewable power storage	-0.5
$P_{g,\max}$(p.u. MW/min), Maximum generation rate	0.12
$P_{g,\min}$(p.u. MW/min), Minimum generation rate	-0.12

III. PROPOSED SOLUTION

A. Neural Actor-Critic

Neural actor-critic architecture is special type of RL agent implies the parallel work of two artificial neural networks: the actor network $\mu(s \mid \theta_\mu)$ and the critic network $Q(s, a \mid \theta_Q)$. In the proposed solution the critic network is an algorithm tracking errors from interaction of actor network a_t with environment s_t according to selected policy. The critic network corrects them in order to find potential right estimation of actor action a'_t, which predicts max possible reward r, see [19] for more technical details. The key advantage of reinforcement learning algorithms is the data-driven study based on interaction with environment. In other words, when agent makes an action, it expects to receive the reward $+r$ or the punishment $-r$. Control mechanism can be briefly summarized as follows. The measured variables Δf, ΔP_L, ΔP_m formulate observation for RL agent. In this paper Δf is considered as the control error and other variables as disturbances. At the same time calculated deviations Δf and ΔP_m go to the block "calculate reward" to reward or punish neural RL actor. General structure of the proposed controller is shown in Fig. 2.

B. Deep Deterministic Gradient Descent

Next, we briefly explain the key application details of the proposed controller optimized by deep deterministic policy gradient (DDPG). DDPG is a model-free reinforcement learning algorithm, designed for tasks with low-dimensional continuous action space [20]. Optimization phase introduces the fusion of deep Q-learning (DQL) and deterministic policy gradient (DPG), which inherits from DQL strategy neural actor-critic. The principle of optimization is based on the search of a minimal difference between target action-value

function of the policy y_i and the critic network $Q(s_i, a_i \mid \theta_Q)$ at each actor network $\mu(s \mid \theta_\mu)$ decision a_t (K_{VI}) per step i in the state s_t (i.e., Δf_t). This is done to minimize the loss function L and receive the max possible reward r_t per training episode. Next, we summarize the overall procedure in the form of a pseudo-code as shown in Algorithm 1.

Algorithm 1 DDPG Algorithm.

1: Initialize critic $Q(s, a \mid \theta_Q)$ and actor $\mu(s \mid \theta_\mu)$ networks with random weights θ_Q and θ_μ.

2: Initialize target network Q' and μ' with $\theta_{Q'} \leftarrow \theta_Q$, $\theta_{\mu'} \leftarrow \theta_\mu$, respectively.

3: Initialize replay buffer R.

4: **for** $episode = 1$ to M **do**

5:     Receive initial process observation as state s_1.

6:     **for** $t = 1$ to T **do**

7:         Select action $a_t = \mu(s_t \mid \theta_\mu)$ according to current policy and disturbances exploration.

8:         Execute action a_t. Observe reward r_t, state s_{t+1}.

9:         Store transition (s_t, a_t, r_t, s_{t+1}) in R.

10:         Sample random minibatch of N transitions (s_i, a_i, r_i, s_{i+1}) from R.

11:         Set $y_i = r_i + \gamma Q'(s_{i+1}, \mu'(s_{i+1} \mid \theta_{\mu'}) \mid \theta_{Q'})$.

12:         Update critic by minimizing loss: $L = \frac{1}{N} \sum_i (y_i - Q(s_i, a_i \mid \theta_Q))^2$.

13:         Update actor policy using sampled gradient:

$$\nabla_{\theta_\mu} J \approx \frac{1}{N} \sum_i \nabla_a Q\left(s, a \mid \theta^Q\right)|_{s=s_i, \alpha=\mu(s_i)}$$
$$\times \nabla_{\theta_\mu} \mu(s \mid \theta^\mu)|_{s_i}.$$

14:         Update the target network: $\theta^{Q'} \leftarrow \tau \theta^Q + (1-\tau)\theta^{Q'}$ and $\theta^{\mu'} \leftarrow \tau \theta^\mu + (1-\tau)\theta^{\mu'}$.

15:     **end for**

16: **end for**

Fig. 2. Structure of the proposed RL controller.

C. RL Agent Reward/Punishment System

In this paper, we redesign reward/punishment strategy to provide more accurate punishment action for RL agent. Reward system is organized to transform the measured frequency deviation to the reward/punishment r_t. Unlike [12] here the reward/punishment system incorporates limitations

of frequency deviation and the power generated by traditional power plant. To provide instructions for Δf regulation, system is organized as follows: initial signal is transformed to the absolute value $|\Delta f|$; if $\Delta f < 0.05$ the RL agent receives the reward, otherwise system does the punishment. In order to provide reasonable rewarding for each action, reward is limited to the range $P_m \in \{0.05, \ldots, 2\}$; however, due to the specific procedure of the controller training the punishment is unlimited and multiplied by 2. To force RL agent to adjust renewable energy generation to traditional power plant the reward/punishment system provides a negative reward -10 if $P_m \notin \{0.05, \ldots, 0.45\}$, otherwise reward becomes 0. If agent's actions collect punishment $r_t < -500$, then the command forced replay is activated. The proposed implementation of reward/punishment system can be summarized as:

$$r_t = \begin{cases} \frac{1}{0.5+|\Delta f|}, & \text{if } |\Delta f| < 0.05, \\ -2|\Delta f|, & \text{if } |\Delta f| > 0.05, \\ 0, & \text{if } \Delta P_m \in \{0.05, \ldots, 0.45\}, \\ -10, & \text{if } \Delta P_m \notin \{0.05, \ldots, 0.45\}. \end{cases} \quad (2)$$

D. Frequency Support Controller

Here we present a modified controller for the frequency support, which combines the virtual inertia control and renewable energy control loops. The proposed architecture is designed to provide more advanced frequency support in the low inertia microgrids. Virtual inertia consists of 3 input system (i.e. discrete time integrator, derivative component and initial input), energy storage system and power limiter as depicted in Fig. 3. It has the inertia power saturation limiter ($P_{inertia,\max}$, $P_{inertia,\min}$), which provides additional robustness for tested algorithms and creates limitations for more realistic simulations. The renewable energy control loop is organized in a similar manner. It has the input ΔP_m, energy storage system and power saturation limiter ($P_{RES,\max}$, $P_{RES,\min}$).
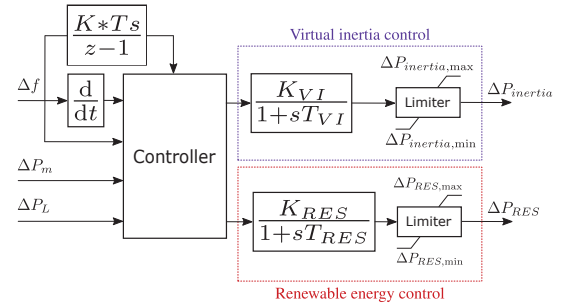
Fig. 3. Structure of frequency support controller.

The main concept used in virtual inertia control is so-called Rate of Change of Frequency (RoCoF), which can be calculated as:

$$RoCoF = \frac{d(\Delta f)}{dt}. \quad (3)$$

The RoCoF defines time derivative of the frequency signal, which is used to calculate the inertia response of as system as:

$$\Delta P_{inertia} = \frac{K_{VI}}{1 + sT_{VI}} \frac{\mathrm{d}(\Delta f)}{\mathrm{d}t}, \quad (4)$$

where T_{VI} is time constant of energy reservation and K_{VI} virtual inertia constant. The renewable energy control flow can be defined as:

$$\Delta P_{RES} = \frac{K_{RES}}{1 + sT_{RES}}(\Delta P_W + \Delta P_{PV}), \quad (5)$$

where

$$\Delta P_W = \frac{1}{1 + sT_{WT}}\Delta P_{wind},$$

$$\Delta P_{PV} = \frac{1}{1 + sT_{PV}}\Delta P_{solar},$$

and T_{RES} is time of renewable energy storage and K_{RES} is renewable energy passing coefficient.

IV. Numeric Results

Consider a microgrid shown in Fig. 1. In this paper we increase the upper limit on a total amount of renewable energy to the level of 50%, and experiment with 100%, 80% and 40% inertia. Figure 4 depicts variation limits of renewable energy $\Delta P_{RES} \in \{0.35, 0.475\}$ and residential loads $\Delta P_L \in \{0.45, 0.75\}$. Microgrid parameters are presented in Table I.
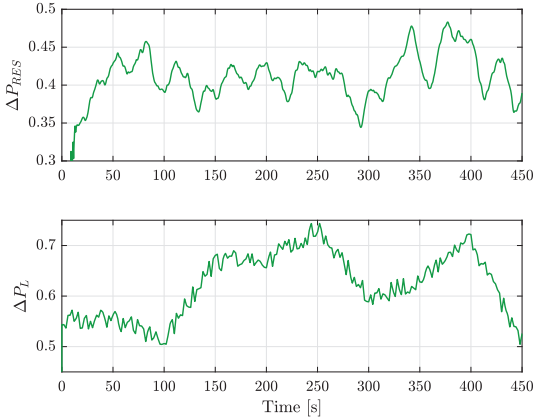
Fig. 4. Variation renewable energy and residential loads during simulations.

In order to validate the proposed method, we compared it with the previously developed single loop algorithm. Simulation results show the reaction of each method on RoCoF variation after contemporary connection of thermal, wind and solar power plants. We consider three different scenarios: Nominal case with high (100%), Scenario I with decreased (80%) and Scenario II with low (40%) system inertia.

Top plot in Fig. 5 shows that without additional learning rules and output, the single loop RL controller cannot provide the effective control strategy and becomes oscillating. The

middle plot shows results similar to those from the nominal case. In fact, according to majority of numerical results the performance of both implementations is affected. The bottom plot shows an observable influence of low inertia on RoCoF change after microgrid being lunched. However, rest of the picture is similar to previous scenarios. Summary of several statistical measures is presented in Table II.
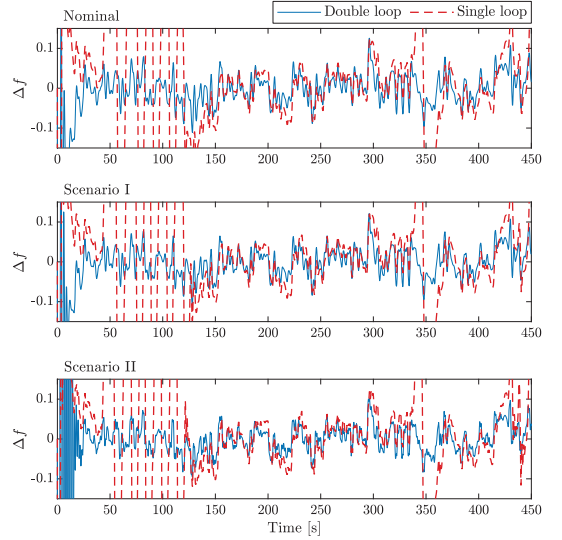
Fig. 5. Frequency variation for inertia coefficients $H = 0.083$ (upper, Nominal), $H = 0.067$ (middle, Scenario I) and $H = 0.032$ (bottom, Scenario II).

TABLE II
SUMMARY OF STATISTICAL MEASURES FOR DIFFERENT LEVELS OF
INERTIA.

Controller	Nominal	Scenario I	Scenario II
Root mean square error (RMSE)			
Double loop RL	0.0359	0.0371	0.0384
Single loop RL	0.2792	0.2791	0.2867
Mean absolute error (MAE)			
Double loop RL	0.0269	0.0269	0.0274
Single loop RL	0.1664	0.1678	0.175
Integral absolute error (IAE)			
Double loop RL	129.74	131.25	136.86
Single loop RL	826.81	828.27	838.58

The proposed double loop RL controller allowed to increase the level of renewable integration energy from 20% to 50% without compromising overall performance. From Table II we can see that the influence of decreased inertia on performance of both control methods is insignificant. At the same time, the increased RES penetration provides a notable affect on performance of the single loop controller. This can be seen in the form of oscillations of signal Δf, which happen shortly

after the microgrid lunch and periodically repeat. It happens due to conflict in power generation between traditional power plant and renewable energy. Proposed modified reinforcement learning controller is designed to avoid such unwanted situations and yields better performance in all scenarios than the previously proposed single loop controller.

V. Conclusion and Discussion

Presented paper demonstrates the potential of extended reinforcement learning based controller in mitigation of low inertia phenomena with high penetration of renewable energy. Reinforcement learning control with single loop and single input reward system cannot provide satisfactory control quality in case of microgrids with high renewable energy integration. Therefore, in this paper we propose modified reward/punishment system and control architecture of reinforcement learning based controller for virtual inertia and renewable energy flow control. The comparative study is based on three different scenarios in conditions with high renewable energy integration. In the future research we plan to apply more advanced models of energy storage, and develop a software for RL based controller stability validation.

References

[1] "Renewables global futures report," REN21, Tech. Rep., 2017.
[2] B. Kroposki, B. Johnson, Y. Zhang, V. Gevorgian, P. Denholm, B.-M. Hodge, and B. Hannegan, "Achieving a 100% renewable grid: Operating electric power systems with extremely high levels of variable renewable energy," *IEEE Power and Energy Magazine*, vol. 15, no. 2, pp. 61–73, 2017.
[3] G. Magdy, E. A. Mohamed, G. Shabib, A. A. Elbaset, and Y. Mitani, "Microgrid dynamic security considering high penetration of renewable energy," *Protection and Control of Modern Power Systems*, vol. 3, no. 1, 2018.
[4] D. E. Olivares, A. Mehrizi-Sani, A. H. Etemadi, C. A. Canizares, R. Iravani, M. Kazerani, A. H. Hajimiragha, O. Gomis-Bellmunt, M. Saeedifard, R. Palma-Behnke, G. A. Jimenez-Estevez, and N. D. Hatziargyriou, "Trends in microgrid control," *IEEE Transactions on Smart Grid*, vol. 5, no. 4, pp. 1905–1919, 2014.
[5] M. Dreidy, H. Mokhlis, and S. Mekhilef, "Inertia response and frequency control techniques for renewable energy sources: A review," *Renewable and Sustainable Energy Reviews*, vol. 69, pp. 144–155, mar 2017.
[6] K. Y. Yap, C. R. Sarimuthu, and J. M.-Y. Lim, "Virtual inertia-based inverters for mitigating frequency instability in grid-connected renewable energy system: A review," *Applied Sciences*, vol. 9, no. 24, p. 5300, 2019.
[7] G. Magdy, A. Bakeer, G. Shabib, A. A. Elbaset, and Y. Mitani, "Decentralized model predictive control strategy of a realistic multi power system automatic generation control," in *Nineteenth International Middle East Power Systems Conference*, 2017.
[8] G. Magdy, G. Shabib, A. A. Elbaset, and Y. Mitani, "A novel coordination scheme of virtual inertia control and digital protection for microgrid dynamic security considering high renewable energy penetration," *IET Renewable Power Generation*, vol. 13, no. 3, pp. 462–474, 2019.
[9] T. Kerdphol, F. S. Rahman, M. Watanabe, and Y. Mitani, "Robust virtual inertia control of a low inertia microgrid considering frequency measurement effects," *IEEE Access*, vol. 7, pp. 57 550–57 560, 2019.
[10] H. Ali, G. Magdy, B. Li, G. Shabib, A. A. Elbaset, D. Xu, and Y. Mitani, "A new frequency control strategy in an islanded microgrid using virtual inertia control-based coefficient diagram method," *IEEE Access*, vol. 7, pp. 16 979–16 990, 2019.
[11] N. Sockeel, J. Gafford, B. Papari, and M. Mazzola, "Virtual inertia emulator-based model predictive control for grid frequency regulation considering high penetration of inverter-based energy storage system," *IEEE Transactions on Sustainable Energy*, vol. 11, no. 4, pp. 2932–2939, 2020.
[12] V. Skiparev, J. Belikov, and E. Petlenkov, "Reinforcement learning based approach for virtual inertia control in microgrids with renewable energy sources," in *IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe)*, The Hague, NL, 2020.
[13] T. Kerdphol, M. Watanabe, K. Hongesombut, and Y. Mitani, "Self-adaptive virtual inertia control-based Fuzzy logic to improve frequency stability of microgrid with high renewable penetration," *IEEE Access*, vol. 7, pp. 76 071–76 083, 2019.
[14] T. Kerdphol, F. Rahman, Y. Mitani, K. Hongesombut, and S. Küfeoğlu, "Virtual inertia control-based model predictive control for microgrid frequency stabilization considering high renewable energy integration," *Sustainability*, vol. 9, no. 5, p. 773, 2017.
[15] W. Guo, F. Liu, J. Si, and S. Mei, "Incorporating approximate dynamic programming-based parameter tuning into PD-type virtual inertia control of DFIGs," in *International Joint Conference on Neural Networks*, 2013.
[16] D. Shrestha, U. Tamrakar, N. Malla, Z. Ni, and R. Tonkoski, "Reduction of energy consumption of virtual synchronous machine using supplementary adaptive dynamic programming," in *IEEE International Conference on Electro Information Technology*, 2016.
[17] A. Ulbig, T. S. Borsche, and G. Andersson, "Impact of low rotational inertia on power system stability and operation," *IFAC Proceedings Volumes*, vol. 47, no. 3, pp. 7290–7297, 2014.
[18] H. Bevrani, T. Ise, and Y. Miura, "Virtual synchronous generators: A survey and new perspectives," *International Journal of Electrical Power & Energy Systems*, vol. 54, pp. 244–254, 2014.
[19] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, 2nd ed. The MIT Press, 2018.
[20] 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, 2019.

Appendix 3

VII

V. Skiparev, R. Machlev, N. R. Chowdhury, Y. Levron, E. Petlenkov, , and J. Belikov. Virtual inertia control methods in islanded microgrids. *Energies*, 14(6):1562, 2021

Review

Virtual Inertia Control Methods in Islanded Microgrids

Vjatseslav Skiparev [1],*[ID]**, Ram Machlev** [2][ID]**, Nilanjan Roy Chowdhury** [3][ID]**, Yoash Levron** [2][ID]**, Eduard Petlenkov** [4][ID] **and Juri Belikov** [1][ID]

1   Department of Software Science, Tallinn University of Technology, 12618 Tallinn, Estonia;
    juri.belikov@taltech.ee
2   The Andrew and Erna Viterbi Faculty of Electrical Engineering, Technion—Israel Institute of Technology,
    Haifa 3200003, Israel; ramm@campus.technion.ac.il (R.M.); yoashl@ee.technion.ac.il (Y.L.)
3   Shamoon College of Engineering, Beer-Sheva 84100, Israel; nilanjan2008@gmail.com
4   Department of Computer Systems, Tallinn University of Technology, 12618 Tallinn, Estonia;
    eduard.petlenkov@taltech.ee
*   Correspondence: vjatseslav.skiparev@taltech.ee

Abstract: Although the deployment and integration of isolated microgrids is gaining widespread support, regulation of microgrid frequency under high penetration levels of renewable sources is still being researched. Among the numerous studies on frequency stability, one key approach is based on integrating an additional loop with virtual inertia control, designed to mimic the behavior of traditional synchronous machines. In this survey, recent works related to virtual inertia control methods in islanded microgrids are reviewed. Based on a contextual analysis of recent papers from the last decade, we attempt to better understand why certain control methods are suitable for different scenarios, the currently open theoretical and numerical challenges, and which control strategies will predominate in the following years. Some of the reviewed methods are the coefficient diagram method, H-infinity-based methods, reinforcement-learning-based methods, practical-swarm-based methods, fuzzy-logic-based methods, and model-predictive controllers.

Keywords: frequency control; islanded microgrid; renewable energy; virtual inertia control

1. Introduction

Renewable energy sources (RESs) are frequently deployed in modern power grids to promote a myriad of environmental and economical benefits. However, the increasing integration of RESs significantly decreases the rotational inertia of the grid, which jeopardizes grid stability and its overall dynamic behavior [1–4]. A central challenge is regulating the grid's frequency under high penetration levels of renewable sources. One approach for addressing this problem is to install fast-reacting storage systems with virtual inertia controllers alongside low-inertia power sources; such controllers have been extensively studied in recent years [5–11]. Each control method has its own benefits and limitations. For instance, classical control paradigms are simple in general but are designed for specific scenarios, whereas data-driven algorithms are flexible and enable online learning. However, these algorithms are numerically complex and require adequate data to operate efficiently. Hybrid control strategies have low numeric complexity, but their convergence is hard to guarantee in most cases. Proposing suitable guidelines for choosing the best algorithm is currently an open question, and this question becomes more important when the microgrid is isolated [11–17].

Microgrids have received increasing attention as a means of integrating distributed generation into the electricity grid [18]. Usually described as confined clusters of loads, storage devices, and small generators, these autonomous networks connect as single entities to the public distribution grid through a point of common coupling (PCC). Figure 1 illustrates a typical microgrid network. Microgrids comprise a variety of technologies: renewable sources, such as photovoltaic and wind generators, are operated alongside

traditional high-inertia synchronous generators, batteries, and fuel-cells. Thus, energy is generated near the loads, enabling the use of small-scale generators that increase reliability and reduce losses over long power lines. The locality of the microgrid network enables the improved management of energy. Generators (and possibly loads) may be controlled by a local energy management system to optimize power flow within the network. The objectives of energy management depend on the mode of operation: islanded or grid-connected [19]. In grid-connected mode, the typical objectives are to minimize the price of energy import at the PCC, to improve power factor at the PCC, and to optimize the voltage profile within the microgrid [20]. In islanded mode, which is addressed in this paper, the main goal of power management is to stabilize the system and preserve high reliability and resiliency in terms of frequency and voltage.
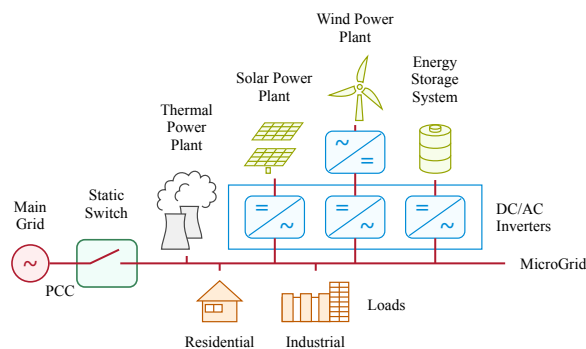
Figure 1. Schematic representation of a typical microgrid. PCC—point of common coupling.

Few recent survey papers describe different aspects in the context of virtual inertia within power grids with a high penetration of RESs. A comprehensive review of virtual inertia implementation techniques was provided in [12]. The reviewed works were classified and compared using virtual inertia topologies. Some selected topologies were simulated, showing that similar inertial responses can be achieved, relating the parameters of these topologies through time and inertia constants. A discussion of the challenges and research directions is presented, indicating future research needs for the integration of virtual inertia systems. Singh et al. [21] reviewed various topologies for emulating a virtual inertia algorithm along with control strategies for general distributed generation. They also reviewed the optimal size and location of synthetic inertia in a power system. Other authors [22] presented a review focusing on the inertia values for power systems. The inertia values were estimated based on different regions in the last 20 years. The contribution of photovoltaic (PV) power plants as virtual inertia was discussed and the damping factor evolution was analyzed.

Contrary to these comprehensive reviews, which focused on virtual inertia topologies implementation [12], virtual inertia and frequency control for distributed energy sources [21], and inertia estimation evolution in power systems [22], we focused on the systematic comparison of virtual inertia control methods designed to solve the frequency regulation problem in islanded microgrids. In particular, we aimed to understand why certain control methods are more efficient in different circumstances, and which control strategies will gain popularity in the coming years. Toward this end, we considered different control techniques available in the literature for the period of 2010–2020, and then categorized them into three groups: classic, advanced, and hybrid methods. We provide a detailed analysis of each control and optimization paradigm through various quality criteria. Finally, we provide a contextual analysis and highlight the current developments and trends for various combinations of virtual inertia control methods and technologies with a focus on microgrid applications.

The rest of this paper is organized as follows: Section 2 presents a model of a standard low-inertia microgrid and explains different control quality criteria. Section 3 summarizes the classical methods applied for virtual inertia control, followed by a discussion of the advanced control methods presented in Section 4. Hybrid control algorithms are described in Section 5. Section 6 provides an analysis of recent trends in low-inertia power systems and virtual inertia control.

2. Overview of Low-Inertia Microgrid System

The low-inertia microgrid encompasses participants with different power generation inertia and loads with complex dynamics [23–26]. Therefore, microgrids with high RES penetration pose various challenges for integration to the massive distribution networks such as (1) active/reactive power imbalance and voltage droop in transmission lines, (2) production/consumption imbalance in distribution loads, and (3) frequency mismatch with other microgrids and the rest of the power grid [3,27]. Hence, energy storage systems are considered the prime actuator in frequency stability control, which, in reality, have physical limitations such as (1) (dis)charge cycles, (2) restricted power reservation, (3) reserved power losses, and (4) individual speed of (dis)charge. Moreover, energy storage control performed by virtual inertia or a virtual synchronous generator (VSG) uses power-inverting electronics, which has delays in frequency measurement and power conversion [12,28–32].

2.1. Modeling of a Low-Inertia Microgrid

The considered microgrid was adopted from several recent publications [16,33–37] and is depicted in Figure 2. The addressed scenario includes simplified residential/industrial loads, energy sources (thermal power plant, wind farm, and solar power plant), and energy storage systems [11,38,39]. The thermal power plant is composed of a governor with a generator rate constraint (GRC) and a turbine with a frequency rate limiter, which restricts the valve opening/closing (V_U and V_L, respectively). The dynamic model of a microgrid uses a hierarchical architecture with primary and secondary control loops. The primary control loop has a droop coefficient $1/R$, and the secondary loop has an area control error (ACE) system with a second frequency controller K_I and a first-order integrator. Frequency regulation is performed by a virtual inertia device with an additional controller. The balancing system is performed as the first-order transfer function with microgrid damping coefficient D and system inertia H, which are common for all generators. The power generation by variable energy sources is modeled as a random signal with a first-order holder. The hierarchical structure includes the reservation of the primary and secondary control loops. The modeling parameters of the microgrid are summarized in Table 1; Table 2 lists the typical simulation scenarios available in the recent literature.

Table 1. Nomenclature: microgrid parameters.

Variable	Physical Meaning
ΔP_m	Generated power change from the distributed generator
T_t	Time constant of the turbine
ΔP_g	Governor valve-position change
T_g	Time constant of the governor
ΔP_{ACE}	Control signal change for secondary control
K_I	Integral control variable gain
ΔP_W	Change in generated power-based wind farm
ΔP_{wind}	Initial wind power variation
T_{WT}	Time constant of wind turbines
ΔP_{PV}	Change in generated power-based solar farm
ΔP_{solar}	Initial solar power variation
T_{PV}	Time constant of the solar system
ΔP_L	Load power change
ΔP_{RL}	Variations in residential loads
ΔP_{IL}	Variations in industrial loads

Figure 2. Schematic representation of an islanded microgrid with hierarchical control. GRC—generator rate constant; LFC—load-frequency control.

2.2. Frequency Regulation in Low-Inertia Power Systems

Frequency stability is important when low-inertia energy sources penetrate the grid in large amounts [1,40,41]. For example, the wind turbine rotor of a synchronous generator has natural inertia, which plays a key role in the power compensation for short periods (up to 5 s) [3]. Solar panels may be considered as zero-inertia generators, since they do not provide physical energy storage [42]. The response of frequency deviation is defined by the rate of change of frequency (RoCoF), which can be calculated as follows [43,44]:

$$\text{RoCoF} = \frac{d(\Delta f)}{dt}. \tag{1}$$

The magnitude of the RoCoF reflects the balanced state in the dynamics of renewable power sources. The problem is generating an active power resembling that generated by traditional power plants.

Table 2. Nomenclature: dynamic parameters of islanded microgrids in different scenarios.

Name	Uncertainty Parameter	Nominal Value	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5
System inertia	H (p.u. MW s)	95–100%	80%	40–50%	25–30%	15%	10%
Droop characteristic	R (Hz/p.u. MW)	2.4	2.4	2.4	1.8–2.4	2.4	1.2
Time constant of governor	T_g (s)	0.1–0.12	0.1	0.1	0.1–0.15	0.1	0.175
Time constant of turbine	T_t (s)	0.4–0.975	0.4	0.4	0.4–0.7	0.4	0.7
Time constant of solar panel	T_{PV} (s)	1.8–1.85	1.85	1.8–1.85	1.8	1.85	1.85
Time constant of wind turbine	T_{WT} (s)	1.5	1.5	1.5	1.5	1.5	1.5
Integral control variable gain	K_i (s)	0.05	0.05	0.05	0.04–0.05	0.05	0.03
System damping coefficient	D (p.u. MW/Hz)	0.015–0.0195	0.015	0.015	0.0135–0.015	0.015	0.003
Frequency bias	β (p.u. MW/Hz)	1.0	1.0	1.0	0.8–1.0	1.0	0.7

<center>**Table 2.** *Cont.*</center>

Name	Uncertainty Parameter	Nominal Value	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5
Virtual inertia control gain	K_{VI}	0.5–0.8	1.0	1.0	0.8–1.0	1.0	0.4
Virtual inertia time constant	T_{VI} (s)	10	10	10	10	10	11
Virtual inertia control power limiter	$\Delta P_{inertia,\max/\min}$	0.25–0.3	0.25-0.3	0.25–0.3	0.3	0.3	0.25
Valve gate speed	$V_{U/L}$	0.3–0.5	0.5	0.1–0.5	0.1–0.5	0.3	0.5
Time constant phased locked-loop (PLL)	ω_n (s)	1.5	–	–	–	–	0.3
References		[16,33,35–37,45]	[34,37]	[16,33,34, 37,45,46]	[16,34]	[33]	[35]

2.3. Virtual Inertia Control

The virtual synchronous generator (VSG) produces the power alternative to the real synchronous machine [47,48]. This generator can be applied in systems with a high level of fluctuating renewable power to enhance the frequency stability. Virtual inertia (VI) is a specific part of the VSG designed to compensate for the lack of inertia using a power injection mechanism [3]. The default operational limitations of the virtual inertia device cannot provide reliable frequency support. Therefore, an additional robust controller must be used to deal with nonlinearities in low-inertia environments. Traditionally, the virtual inertia control setup (Figure 3) consists of a derivative component, a designed controller $K(s)$, virtual inertia control (energy storage system and virtual inertia variable gain), and a power limiter ($\Delta P_{inertia,\max}$, $\Delta P_{inertia,\min}$).
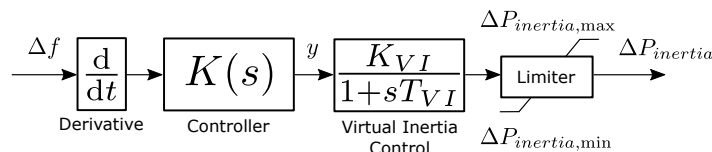
Figure 3. Typical structure of a virtual inertia control block.

2.4. Energy Storage System

The energy storage system (ESS) has been implemented in various physical realizations [38,49,50]. The technology can be directly incorporated into frequency-response services and support the RoCoF during a frequency event. For the last decade, ESSs became an essential component in renewable energy integration, since they may provide frequency smoothness and balance for further dispatch [5,51–57]. The simplified ESS model can be represented as follows:

$$G(s) = \frac{1}{T_{VI}s + 1}. \tag{2}$$

2.5. Hierarchical Control

Hierarchical frequency control introduces a multilevel cascade system with three key layers: primary, secondary (load frequency control), and tertiary control, and two additional layers: internal generation control and high-level policy control. Primary control is responsible for regulation of individual elements: power sharing, frequency droop, and local voltage control. Secondary control is oriented toward the balancing of active and reactive power by determination of the set-points of the primary controller and secondary control, including grid synchronization, automatic generation control (AGC), secondary load-frequency control (LFC), and voltage-drop control. Tertiary control (i.e., reserved) is

related to energy management. It is used to provide optimal power flow and steady-state conditions in a distribution network [3,58–61].

2.6. Control Quality Criteria

In this section, we discuss typical control criteria. They are then used to examine the benefits and drawbacks of the presented algorithms.

- Online learning. Real-time optimization is used to adapt controllers to varying conditions [62,63], and usually requires a special computational device for additional tuning, validation, and verification.
- Robustness. Flexible reaction to disturbances is an important requirement in low-inertia power grids, since, in practice, frequency deviation is limited to the range of ± 0.1–1.5 Hz [12,64–68], and Nadir requirements are ± 0.024 Hz [12,69]. Using this criterion, we briefly describe testing scenarios and the performance of the considered virtual inertia controller.
- Implementation complexity. Complexity corresponds to the implementation efforts of an algorithm in real controlling hardware: size, number of inputs and outputs, mathematical complexity, etc.
- Optimization difficulty. This depends on the number of inputs and outputs, time for optimization procedure, and other requirements for the computational power to provide the maximum possible efficiency.

2.7. Description of the Virtual Inertia Control Algorithms

Several recent works [12,16,33–36,70,71] addressed the problem of optimal frequency support with high penetration of variable renewable energy sources. For example, Kerdphol et al. [35] designed a robust H_∞ controller to provide stability support based on the rate of change of frequency. The proposed solution provides advantages over conventional virtual inertia control and optimally tuned proportional integral (PI) controllers in scenarios when the wind farm is connected, solar panels are disconnected, and the system inertia is 10% and 100%, respectively. Kerdphol et al. [34] further studied the problem by implementing a virtual inertia control scheme combined with a fuzzy-logic-based approach. The proposed algorithm performed robustly under different scenarios with additional uncertainties, including 80%, 40%, and 30% total system inertia and mismatches in the primary/secondary control loops. Kerdphol et al. [45] proposed a model predictive control scheme and compared it to a fuzzy-logic controller for the case of additional load connections. Unlike the previous works, the studied microgrid has conceptual differences: a closed-loop turbine system, RES power generation from two complex wind farms, and minor differences in the transfer function describing the turbine and system inertia. Similar ideas were presented in Tamrakar et al. [72], but without modeling renewable energy disturbances. Magdy et al. [16] presented a PI controller optimized using particle swarm optimization and combined with a digital frequency protection system in scenarios of (dis)connecting loads and renewable energy sources.

In the following sections, we discuss the main features and constructive advantages and disadvantages of the most common algorithms for virtual inertia control, focusing on the load-frequency stability, implementation complexity, and performance against disturbances. We categorize the revised implementations into three groups: advanced, classical, and hybrid control as detailed in Figure 4.

Advanced	**Hybrid**	**Classical**
Reinforcement Learning	Model Predictive Control	Coefficient Diagram Method
Fuzzy-Logic Control	Particle Swarm Optimization + PI Controler	H-infinity

Figure 4. Classification of algorithms for virtual inertia control. PI—proportional integral.

3. Classical Control Algorithms

The major features of classical control algorithms are as follows:

- Classical optimization. The optimization is based on the reaction to disturbances, which are approximated by a transfer function or state-space representation. Usually, classical optimization are applied to slow processes [73].
- Simplicity. Classical algorithms have a simple control structure, which enables effective manual tuning and requires low computational power.
- High robustness. Classical algorithms can be highly robust, but require a specific design procedure.

3.1. H-Infinity

H-infinity, H_∞, achieves the synthesis of an optimal controller by considering microgrid disturbances and uncertainties via state-space representation, which can provide high robustness and simple hardware realization. However, the main difficulty is the necessity of designing an accurate state-space description for tuning the controller [33,35]. Frequency control based on H_∞ was used in [33,35,74–76]. The solution presented in [35] applies a linear fractional transformation in the optimal H_∞ regulator design as the basis for modeling microgrid uncertainties z, such as system inertia H, damping properties D, and phased locked-loop (PLL) delays (ω_n and ζ).

H_∞ optimization performs in offline mode and is more vulnerable to low-inertia nonlinearities than data-driven algorithms. At the same time, synthesis of the robust model by H_∞ provides reliable frequency support. For example, Kerdphol et al. [33] implemented this method, which was successfully tested with 95%, 45%, and 15% of the nominal system inertia and using two types of disturbances: (1) 10% of step changes in load power demand and (2) mismatch in microgrid generation by increased time constant of the governor and time constant for the turbine. Kerdphol et al. [35] reported an H_∞ controller tested with 100% and 10% system inertia in a scenario with 80% renewable energy penetration. However, the common limitation of the H_∞ method is the notable peaks during (dis)connection of power plants. H_∞ requires a detailed understanding of classical control theory and optimization, which does not require powerful hardware for operation. Nevertheless, the synthesized control model is a high-order transfer function and often requires order reduction [33,35]. The biggest difficulties with H_∞ optimization are the procedure for developing an accurate state-space representation and the manual estimation of disturbances. The optimization based on application of the H_∞ controller is summarized in Algorithm 1.

Algorithm 1 Design of the H_∞ controller

1: Define the state vector $x^\mathsf{T} = [\Delta f, \Delta P_m, \Delta P_g, \Delta P_{ACE}, \Delta P_{inertia}, \Delta P_W, \Delta P_{PV}, \Delta f_{PLL1}, \Delta f_{PLL2}]^\mathsf{T}$
2: Define the distribute vector $w^\mathsf{T} = [\Delta P_{wind}, \Delta P_{solar}, \Delta P_L]^\mathsf{T}$
3: Define the control input $u = \Delta f_{PLL}$ and output $y = \Delta f_{PLL} K(s)$
4: For a given microgrid, derive the state-space model with defined vectors
5: Design the optimal H_∞ controller using the linear fractional transformation technique
6: Validate the designed $K(s)$ controller using a close-loop inequality equation, and if needed, repeat the optimization procedure

---

3.2. Coefficient Diagram Method

Controllers based on the coefficient diagram method (CDM) rely on an algebraic optimization approach through polynomial state-space representation and the Routh–Hurwitz stability criterion [77,78], where the theoretical basis is constituted to satisfy the Lipatov–Sokolov stability criterion [79–82].

Similar to H_∞, the optimization procedure is designed for offline mode. The implementation of a controller based on the CDM in Ali et al. [36] produced frequency stability in a range less than ± 0.1 Hz in a scenario with 100% inertia and two types of disturbances: (1) 10% step load perturbation and (2) random load demand. In contrast to H_∞, it can mitigate peaks after the (dis)connection of renewable energy sources. However, the solution uses a two degrees of freedom system structure expressed as $N(s)/D(s)$, which is designed to track a limited number of disturbances. The main drawback of CDM controller synthesis is similar to H_∞: it relies on a good understanding of classical control theory optimization. However, it can be implemented using relatively simple hardware. In contrast to H_∞, CDM optimization performs without requiring order reduction and uses the coefficient method instead of the Bode diagram [36,83]. However, the validation of synthesized control by Routh–Hurwitz or Lipatov–Sokolov stability criteria depends on the order of the synthesized control system [36,81–83]. The design procedure is summarized in Algorithm 2, which was adopted from the flowchart provided by [36].

Algorithm 2 CDM algorithm

1: Define polynomial equation for microgrid modeling
2: Define external disturbances $d = [P_{wind}, P_{solar}, P_L]$ and reference input $r = \Delta f_{ref}$ and CDM controller as $K(s)$
3: Calculate the $K(s)$ output system as $y = \Delta P_{inertia}$ and the input system as $u = \Delta f$ with external disturbances and reference input
4: Calculate the polynomial of the designed $K(s)$ control system with microgrid external disturbances
5: Validate the designed $K(s)$ control system
6: **if** the stability conditions of optimal CDM controller are verified **then**, go to step 10
7: **else** Check the value of the stability indices and the stability limits
8:     Calculate the desirable CDM controller $K_{target}(s)$
9:     Compare $K_{target}(s)$ and $K(s)$,
10:     **if** the model is validated **then**
11:         A robust $K(s)$ controller is obtained
12:         Check the robustness of the system response
13:     **else** Repeat the procedure
14:     **end if**
15: **end if**

---

4. Advanced Control Algorithms

The major features of advanced control algorithms can be expressed as follows:

- Adaptation to uncertain conditions. Advanced control algorithms may provide adaptive reactions to disturbances that were not predicted.

- Prediction-based optimization. Fast processes, such as electrical frequency variation, are easier to predict than postreaction. This principle gives additional advantages, because data-based optimization follows the events prediction model. The drawback of the approach is the necessity to design a memory buffer for data recording and further prediction-based tuning.
- Online learning. Data-driven optimization implies recorded data analysis of controlled processes. When conditions are changing radically, this approach provides a strategy for optimization of controller parameters in parallel with real-time control.
- Complexity. Advanced algorithms require a powerful computing system. The main benefit from complexity is effective multiloop control and adaptation to process dynamics.

4.1. Reinforcement Learning-Based Controller

Reinforcement learning (RL) is an agent-based and model-free machine learning algorithm [84]. The main approach of RL optimization is based on trial and error, which allows direct validation of the artificial neural network (ANN)-based controller with the control object and prediction of negative consequences [37,84–86]. The benefit of this method is mandatory data-driven optimization, which is naturally designed for online learning. In [37], RL was compared with H_∞, producing slightly better performance in terms of frequency stability in scenarios with 100%, 80%, and 40% inertia and connection of wind, solar, and thermal plants during the launch of industrial and residential loads, and 20% RES penetration. Since the algorithm uses a deep neural network, it requires strong computational hardware and is relatively complex for implementation. The method requires selection of an optimal action $a(t)^*$ at each step $s(t)$ and takes a long time. For RL, it is necessary to design a proper reward system and to choose the right training strategy, which may differ [37,87–89]. For example, in previous works [87,88], the RL optimization for frequency support was performed by approximated dynamic programming. In contrast, Skiparev et al. [37] used the deep deterministic policy gradient to train an RL-based controller for virtual inertia emulation. The optimization mechanism using the RL algorithm is summarized in Algorithm 3.

Algorithm 3 Reinforcement-learning-based algorithm

1: Define the actor and critic neural networks
2: Define $a_t = \Delta P_{inertia}$, $s_t = \Delta f_t$, and $s_{t+1} = \Delta f_{t+1}$ of RL controller
3: Define the desirable total reward for the RL controller r_{target}
4: Start training the RL-based controller
5: Receive initial process observation of microgrid dynamics as state s_1
6: Select action a_t of the actor network according to current policy and disturbances exploration
7: Execute action a_t of the actor network
8: Observe reward r_t and state s_{t+1} using the critic network
9: **if** $r \geq r_{target}$ **then** Controller training successfully completed
10: **else** Continue training
11: **end if**

4.2. Fuzzy Logic Controller

Fuzzy logic controller design provides effective manual optimization compared with other advanced algorithms. Several examples of frequency regulation can be found in the literature [34,62,90–94]. Since fuzzy-logic-based controllers can be manually tuned, the data-driven approach is optional. Correct configuration of the controller can create a robust system. Kerdphol et al. [34] applied a standard fuzzy logic controller for virtual inertia control, which was capable most of the time of holding Δf inside the ± 0.1 Hz band with 80%, 60%, and 30% system inertia in scenarios with 20% and 80% RES penetration and mismatch in primary/secondary control loops. Controller design requires a good understanding of fuzzy rules design principles. In addition, the method requires powerful

hardware for implementation. However, it uses fuzzy logic without an optimizer, which can be considered a drawback, since it requires the manual design of the optimal fuzzy rules [92,95].

5. Hybrid Control Algorithms

Hybrid algorithms inherit features from both categories. Model predictive control (MPC) is an example of a controller that cannot be classified into either of the above-mentioned categories. Optimization can be based on state-space representation [45] or input/output (I/O) relation approximated by the data-driven approach [96]. The PI controller optimized by particle swarm optimization (PSO) is another hybrid example, combining a simple controller with the data-driven approach [16].

5.1. Evolutionary Optimization

Particle swarm optimization is a popular evolutionary algorithm inspired by collective species behavior such as flocks of birds [97]; stochastic optimization should provide the best performance through searching for a global minima. The particle swarm strategy is a stochastic data-driven optimizer that enables online learning [16,56,98,99]. Magdy et al. [16] used PSO for optimal tuning of a PI controller via searching the global minima of a microgrid, which provided robust control with 100%, 80%, and 30% system inertia. The performance of the optimal PI in Magdy et al. [16] showed relatively stable frequency support with 100%, 50%, and 30% system inertia and with 57% RES penetration. In contrast with other solutions, Magdy et al. [16] applied a dynamic model of a microgrid with digital protection, which provided additional frequency stability. PI/PID is a widely used controller in the power industry due to its simple construction [100–102]. However, the PSO algorithm is a self-learning optimizer, which is more complex for implementation. To produce an optimally tuned PI controller, the optimizer has to consider the state-space dynamic modeling of microgrid uncertainties, which requires a relatively long time to find optimal settings. The PSO procedure is summarized in Algorithm 4, adopted from [16].

Algorithm 4 PSO algorithm

1: Define microgrid state-space matrix
2: Define state vector $X^\mathsf{T} = [\Delta f, \Delta P_g, \Delta P_m, \Delta P_{WT}, \Delta P_{PV}, \Delta P_{inertia}]^\mathsf{T}$
3: Define external disturbances vector $W^\mathsf{T} = [\Delta P_{Wind}, \Delta P_{Solar}, \Delta P_L]^\mathsf{T}$
4: Define the control output signal as $Y = [\Delta f]$
5: Compute the state-space model for a given microgrid with defined inputs and outputs
6: Initialize the D-dimension of particles as PI/PID controller coefficients
7: Perform optimization by minimization of the fitness function for each particle
8: Calculate the velocity and current position of each particle. Validate the optimized PI/PID controller
9: **if** stopping criteria of PI/PID controller are met **then**
10:    Optimal parameters of PI/PID are obtained
11: **else** Repeat optimization
12: **end if**

5.2. Model Predictive Control

The model predictive controller (MPC) requires the development of a robust prediction model based on a detailed representation of the process dynamics via collected data [45,103,104]. As a hybrid algorithm, the MPC can be implemented with data-driven [105] or finite-time-horizon [46,106] optimization approaches. Kerdphol et al. [45] applied finite impulse response optimization for model prediction based on the virtual inertia emulation with microgrid state-space representation.

Regarding optimization, MPC can provide real-time learning through data-driven and finite-horizon approaches. According to Kerdphol et al. [45], MPC performance is higher than that of the fuzzy-logic-based controller, and may provide better Δf stability

during (1) (dis)connection of RES power, (2) sudden load change, and (3) mismatch in the main thermal generation scenarios with 100%, 50%, and 25% system inertia and 34% RES penetration. Implementation of the model-prediction-based controller depends on the type of prediction model. The controller requires the calculation of each time sample and heavily depends on the designed model used in the predictions of microgrid disturbances [45,72]. Specifically, Kerdphol et al. [45] used the finite impulse response, which considers each sampling instant in the prediction of microgrid disturbances. The general concept of MPC optimization is summarized in Algorithm 5.

Algorithm 5 MPC algorithm

1: Define the MPC controller as $K(s)$
2: Define the MPC controller input as $u = \Delta f$, output as $y = K(s)\Delta f$, and the desired profile as $r = \Delta f_{ref}$
3: Predict the microgrid dynamics for the current time
4: Optimize the first control step of $K(s)$
5: Adjust the first control step according to MPC control rules
6: Implement the local MPC controller
7: **if** Evaluate the disagreement of tracking consensus with constrains **then**
8:     End MPC optimization
9: **else** Repeat optimization
10: **end if**

---

6. Recent Directions and Trends

One goal of this study was to highlight the popularity of various control methods for virtual inertia emulation reflected in the recent literature. Such trends are explored in this section based on the contextual analysis of additional virtual inertia control. Based on this analysis, we explain the motivation for the choice of several optimal control methods and try to better understand why and when the reviewed methods are most efficient. Special attention is paid to the analysis of relevant keywords describing each method and application area. The fuzzy logic controller, model predictive control, coefficient diagram method, and H-infinity methods are well-defined by their names. However, reinforcement-learning and evolution algorithms are often defined by a specific strategy. Therefore, we used several of the most common types of these optimizations during our literature search. The keywords we used for the control methods are summarized in Table 3. The search was also restricted to the title, abstract, and keywords fields.

Table 3. Search expressions that were used in the literature search.

Primary Expression	Secondary Expression	Third Expression
"virtual inertia control"	"microgrid"	"FLC OR Fuzzy Logic Controller" "MPC OR Model predictive control" "PI/PID" "EA OR GA OR Evolution algorithms OR Genetic algorithms" "CDM OR Coefficient diagram method" ''H_∞ OR H-infinity'' "PSO OR Particle swarm optimization " "RL OR Reinforcement learning "

Figure 5 depicts the rising trend in publications on virtual inertia control over an 11-year period. The Scopus database produced 404 papers and IEEE Xplore produced 239 papers.
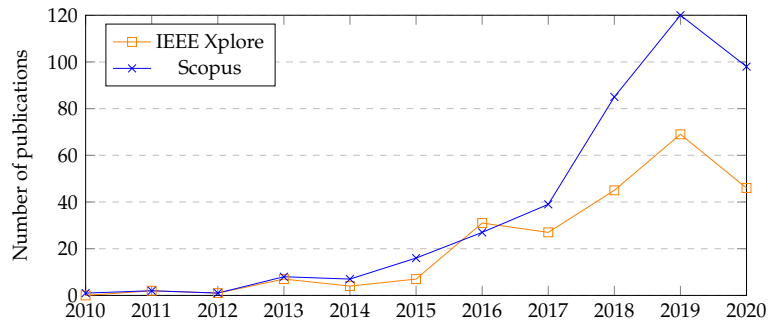
Figure 5. Yearly number of publications in the period from 2010 to 2020 on virtual inertia control.

The frequency-support-related algorithms mostly continued the rising trend, as detailed in Figure 6. To provide a more in-depth analysis, we selected several algorithms commonly used in frequency-control applications. Fuzzy logic and PI/PID appeared to be the most popular control algorithms. Publications indicate the stable interest in usage of PID controller, which can be further equipped with an additional optimization loop based on data-driven algorithms and/or combined with advanced controllers [16,62,99,107,108]. Due to the natural ability in finding global minima, evolution algorithms (e.g., PSO, firefly, and bat) are mostly combined with the fuzzy logic controller (FLC) and/or PID [16,62,99] as one of the most frequently used hybrid algorithms of the existing control loops. Model predictive controllers gained similar attention; in recent years, they have become the most popular. One notable rise was found in the usage of the reinforcement-learning-based strategies, which may become even more popular in the next years due to their ability to perform effective study based on interactions with the environment [37,85,88]. Therefore, we think that the data-driven algorithms will attract more attention in the coming years due to the growing prevalence of data mining and cloud technologies.
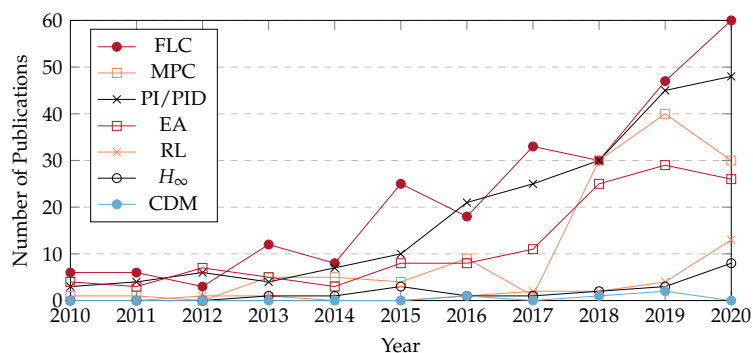
Figure 6. Trends in the frequency-support algorithms in microgrids. FLC—fuzzy logic controller, MPC—model predictive control, EA—evolution algorithm, RL—reinforcement learning, H_∞—H-infinity, CDM—coefficient diagram method.

Figure 7 depicts the search results for the specific technologies used for frequency regulation in microgrids. Energy storage appears to be the most widely used technology. Virtual synchronous generators, virtual inertia, and phase locked-loop have small numbers of publications, since each technology related to synthetic inertia generation is individual

and requires specific design and case studies. Notably, many possibilities exist for research into VSG/VI-related applications [12,22].
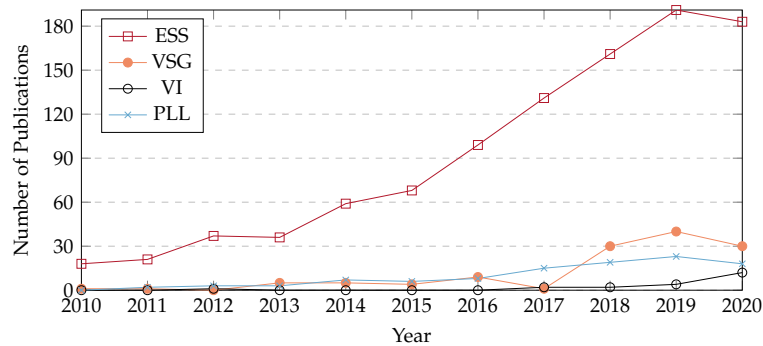
Figure 7. Trends in frequency-control technologies in microgrids. ESS—energy storage system, VSG—virtual synchronous generator, VI—virtual inertia, PLL—phased locked-loop.

Based on analysis of the above trends, it is reasonable to conjecture that in the coming years, the virtual inertia problem will remain in the focus of the community. The isolated microgrid, as a part of the general power grid, faces several important challenges such as active and reactive power balance, power losses in transmission lines, grid frequency out-matching, power production/consumption balance, among others [109]. Most microgrids use simplified models of domestic loads, power plants, and energy storage systems. The European Commission reported the potential research challenges in the renewable energy area in the period of 2021–2027:

- Integrated local energy systems, microgrids, and modular solutions [110–113];
- Cross-border cooperation in transmission grids [110,114–117];
- Electrical transport (cars, trucks, ships, etc.) [110,118–120];
- Effective energy management in domestic appliances (HVAC, boilers) by demand-side management technologies [110,113,115,121,122];
- Solutions for the integration of energy systems and coupling of different energy vectors, networks, and infrastructures in the context of a digitalized, green, and cybersecure energy system [110,113,123].

According to the REN21 report, 63% of world experts agree that by 2050, power generation will focus on centralized or decentralized renewable energy [119] and 71% agree that the transition to 100% renewable energy on a global level is feasible and realistic [119]. In addition, most experts agree that renewable energy should provide at least 32% of the EU energy consumption by 2030 [2,119]. Hence, there is a clear need for continuing the research on and adoption of various solutions, supporting the integration of renewable energy sources; microgrids will most likely play a key role in achieving these goals.

7. Conclusions

Here, we reviewed recent works related to virtual inertia control methods designed to solve the frequency regulation problem in islanded microgrids, with an attempt to better understand the unique characteristics, common uses, and mathematical foundations of the most popular control methods. The control techniques on which we chose to focus were selected following an in-depth content analysis of various sources from the main databases, as detailed in Section 6. This analysis revealed interesting trends in the current research, and may help to understand why certain control methods are more efficient in different circumstances (Table 4), and which control strategies will gain popularity in the coming years.

Table 4. Comparison of virtual inertia control algorithms: advantages, drawbacks, and quality criteria.

	Hand. App.	Online/ Offline	Advantages	Drawbacks	Computational Complexity	Robustness	Optimization Complexity	Refs.
Classical algorithms	Robust H-infinity	Offline	• Robust frequency control • Strong overshoot minimization	• Significant peaks during connection disturbances • Need for order reduction • Limited robustness	Medium	High	Medium	[36,74,75,124–126]
	Coefficient diagram method	Offline	• Higher robustness • No need for order reduction	• Limited robustness	Medium	High	Medium	[36,81,83,127]
Advanced algorithms	Fuzzy-logic-based controller	Online	• Flexible reaction	• Limited by fuzzy rules adaptation • Manual optimization • Long computational time	High	High	High	[16,34,45,128]
	Reinforcement-learning-based controller	Online	• Reward learning system • Advanced feedback from system • High robustness	• Available sample data are needed • Specific to the reward/punishment optimization	Very High	High	Very High	[37,85,87,88,129]
Hybrid algorithms	PI/PID and particle swarm optimization	Online	• Low numeric complexity • Simple controller	• Convergence to global optimal solution is not guaranteed • Limited robustness	Low	Low	Low	[16,56,98]
	Model predictive control	Online	• High robustness • Fast reaction based on prediction • Fast optimization	• Need data reserved for prediction model • Complex optimization	High	High	High	[45,46,70,72,106,130]

For instance, the data show that evolutionary algorithms methods are widely used for tuned PI/PID controllers probably since this enables the analysis of stochastic scenarios with nonlinear constraints. However, evolutionary algorithms may converge to local minima and are therefore not suitable for every application. In such cases, classical control methods seem to be the natural choice since they provide simple and effective solutions to the virtual inertia problems whenever grid dynamics are well-defined. If there is uncertainty in the grid dynamic and nonlinear constraints, fuzzy-logic-based controllers are used extensively, although they are limited to specific and manually defined rules; in cases with a large number of rules, the needed resources increase significantly. The controllers based on the coefficient diagram method principle seem to be the least popular method, maybe due to their limitation of tracking only a limited number of disturbances. Artificial neural networks are also increasing in use due to the increasing amounts of available data; specifically, reinforcement-learning methods are commonly used for solving complex problems when a fully satisfactory algorithm is lacking. In our opinion, these trends may change in the near future due to global initiatives related to the integration of electric vehicles into microgrids and due to the continuing integration of renewable energy sources and beyond-the-meter technologies, which may lead to more available data and thus favor the use of new and more efficient controllers with a focus on data-driven approaches.

Concerning future research, since microgrids are increasingly decentralized and less regulated by governments, it is often impractical to study them from the perspective

of one single entity with unlimited information and control span. Therefore, the recent increasing trend in studies of virtual inertia control for isolated microgrids will likely continue. Whereas classic control techniques are still mainly the focus of the community, the wide adoption and integration of technological innovations such as the Internet of things (IoT), cloud technologies, and data processing powers will likely start shifting the main attention toward data-driven control techniques in the coming years. Another topic of interest may be combining virtual inertia control with suitable energy storage as a supportive technological solution in isolated microgrids. To answer this challenge, the development of new optimal control methods can be considered a possible avenue for future research.

Abbreviations

The following abbreviations are used in this manuscript:

ANN	Artificial neural network
EA	Evolution algorithms
ESS	Energy storage system
CDM	Coefficient diagram method
FLC	Fuzzy logic controller
GRC	Generator rate constraint
HVAC	Heating, ventilation, and air conditioning
LFC	Load-frequency control
MPC	Model predictive control
PCC	Point of common coupling
PID	Proportional-integral-derivative
PLL	Phased locked-loop
PSO	Particle swarm optimization
RL	Reinforcement learning
RES	Renewable energy sources
RoCoF	Rate of change of frequency
VI	Virtual inertia
VSG	Virtual synchronous generator

References

1. Ulbig, A.; Borsche, T.S.; Andersson, G. Impact of low rotational inertia on power system stability and operation. *IFAC Proc. Vol.* **2014**, *47*, 7290–7297. [CrossRef]
2. Kroposki, B.; Johnson, B.; Zhang, Y.; Gevorgian, V.; Denholm, P.; Hodge, B.M.; Hannegan, B. Achieving a 100% renewable grid: Operating electric power systems with extremely high levels of variable renewable energy. *IEEE Power Energy Mag.* **2017**, *15*, 61–73. [CrossRef]
3. Milano, F.; Dörfler, F.; Hug, G.; Hill, D.J.; Verbic, G. Foundations and Challenges of Low-Inertia Systems (Invited Paper). In Proceedings of the 2018 Power Systems Computation Conference (PSCC), Dublin, Ireland, 11–15 June 2018. [CrossRef]
4. Kumar, D.; Mukherjee, B.K.; Mathur, H.D.; Siguerdidjane, H.; Bhanot, S. Forecast-based modeling and robust frequency control of standalone microgrids considering high penetration of renewable sources. *Int. Trans. Electr. Energy Syst.* **2020**, *31*. [CrossRef]
5. Boicea, V.A. Energy Storage Technologies: The Past and the Present. *Proc. IEEE* **2014**, *102*, 1777–1794. [CrossRef]

6. Groß, D.; Bolognani, S.; Poolla, B.K.; Dörfler, F. Increasing the Resilience of Low-inertia Power Systems by Virtual Inertia and Damping. In Proceedings of the 10th Bulk Power Systems Dynamics and Control Symposium (IREP 2017), Espinho, Portugal, 27 August–1 September 2017. [CrossRef]
7. Vorobev, P.; Huang, P.H.; Hosani, M.A.; Kirtley, J.L.; Turitsyn, K. A framework for development of universal rules for microgrids stability and control. In Proceedings of the 2017 IEEE 56th Annual Conference on Decision and Control (CDC), Melbourne, VIC, Australia, 12–15 December 2017. [CrossRef]
8. Groß, D.; Dörfler, F. On the steady-state behavior of low-inertia power systems. *IFAC-PapersOnLine* **2017**, *50*, 10735–10741. [CrossRef]
9. Ojo, Y.; Watson, J.; Lestas, I. A Review of Reduced-Order Models for Microgrids: Simplifications vs Accuracy. *arXiv* **2020**, arXiv:2003.04923.
10. Abubakr, H.; Mohamed, T.H.; Hussein, M.M.; Guerrero, J.M.; Agundis-Tinajero, G. Adaptive frequency regulation strategy in multi-area microgrids including renewable energy and electric vehicles supported by virtual inertia. *Int. J. Electr. Power Energy Syst.* **2021**, *129*, 106814. [CrossRef]
11. Kerdphol, T.; Rahman, F.S.; Watanabe, M.; Mitani, Y. *Virtual Inertia Synthesis and Control*; Springer International Publishing: Cham, Switzerland, 2021. [CrossRef]
12. Tamrakar, U.; Shrestha, D.; Maharjan, M.; Bhattarai, B.; Hansen, T.; Tonkoski, R. Virtual Inertia: Current Trends and Future Directions. *Appl. Sci.* **2017**, *7*, 654. [CrossRef]
13. Moon, H.J.; Chang, J.W.; Lee, S.Y.; Moon, S.I. Autonomous active power management in isolated microgrid based on proportional and droop control. *Energy Procedia* **2018**, *153*, 48–55. [CrossRef]
14. Zhu, J.; Zheng, Y.; Wang, Y.; Yuan, Y. Isolated Microgrid Capacity Configuration Considering Economic Risk of Customer Interruption. In Proceedings of the 2019 IEEE Innovative Smart Grid Technologie -Asia (ISGT Asia), Chengdu, China, 21–24 May 2019. [CrossRef]
15. Yap, K.Y.; Sarimuthu, C.R.; Lim, J.M.Y. Virtual Inertia-Based Inverters for Mitigating Frequency Instability in Grid-Connected Renewable Energy System: A Review. *Appl. Sci.* **2019**, *9*, 5300. [CrossRef]
16. Magdy, G.; Shabib, G.; Elbaset, A.A.; Mitani, Y. A Novel Coordination Scheme of Virtual Inertia Control and Digital Protection for Microgrid Dynamic Security Considering High Renewable Energy Penetration. *IET Renew. Power Gener.* **2019**, *13*, 462–474. [CrossRef]
17. Padmawansa, N.U.; Arachchige, L.N.W. Improving Transient Stability of an Islanded Microgrid Using PV Based Virtual Synchronous Machines. In Proceedings of the 2020 Moratuwa Engineering Research Conference (MERCon), Moratuwa, Sri Lanka, 28–30 July 2020. [CrossRef]
18. Vasquez, J.C.; Miret, J.M.G.J.; Castilla, M.; de Vicuña, L.G. Hierarchical Control of Intelligent Microgrids. *IEEE Ind. Electron. Mag.* **2010**, *4*, 23–29. [CrossRef]
19. Akinyele, D.; Belikov, J.; Levron, Y. Challenges of Microgrids in Remote Communities: A STEEP Model Application. *Energies* **2018**, *2*, 432. [CrossRef]
20. Katiraei, F.; Iravani, R.; Hatziargyriou, N.; Dimeas, A. Microgrids management. *IEEE Power Energy Mag.* **2008**, *6*, 54–65. [CrossRef]
21. Singh, D.; Seethalekshmi, K. A Review on Various Virtual Inertia Techniques for Distributed Generation. In Proceedings of the 2020 International Conference on Electrical and Electronics Engineering (ICE3), Gorakhpur, India, 14–15 February 2020. [CrossRef]
22. Fernández-Guillamón, A.; Gómez-Lázaro, E.; Muljadi, E.; Molina-García, Á. Power systems with high renewable energy sources: A review of inertia and frequency control strategies over time. *Renew. Sustain. Energy Rev.* **2019**, *115*, 109369. [CrossRef]
23. Hiskens, I.A.; Fleming, E.M. Control of inverter-connected sources in autonomous microgrids. In Proceedings of the 2008 American Control Conference, Seattle, WA, USA, 11–13 June 2008. [CrossRef]
24. Schiffer, J.; Zonetti, D.; Ortega, R.; Stanković, A.M.; Sezi, T.; Raisch, J. A survey on modeling of microgrids—From fundamental physics to phasors and voltage sources. *Automatica* **2016**, *74*, 135–150. [CrossRef]
25. Guo, F.; Wang, L.; Wen, C.; Zhang, D.; Xu, Q. Distributed Voltage Restoration and Current Sharing Control in Islanded DC Microgrid Systems Without Continuous Communication. *IEEE Trans. Ind. Electron.* **2020**, *67*, 3043–3053. [CrossRef]
26. Jung, S.; Yoon, Y.T.; Huh, J.H. An Efficient Micro Grid Optimization Theory. *Mathematics* **2020**, *8*, 560. [CrossRef]
27. Mendieta, W.; Canizares, C.A. Primary Frequency Control in Isolated Microgrids Using Thermostatically Controllable Loads. *IEEE Trans. Smart Grid* **2021**, *12*, 93–105. [CrossRef]
28. Van, T.V.; Visscher, K.; Diaz, J.; Karapanos, V.; Woyte, A.; Albu, M.; Bozelie, J.; Loix, T.; Federenciuc, D. Virtual synchronous generator: An element of future grids. In Proceedings of the 2010 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT Europe), Gothenburg, Sweden, 11–13 October 2010. [CrossRef]
29. Zhong, Q.C.; Weiss, G. Synchronverters: Inverters That Mimic Synchronous Generators. *IEEE Trans. Ind. Electron.* **2011**, *58*, 1259–1267. [CrossRef]
30. Bevrani, H.; Ise, T.; Miura, Y. Virtual synchronous generators: A survey and new perspectives. *Int. J. Electr. Power Energy Syst.* **2014**, *54*, 244–254. [CrossRef]
31. Zhong, Q.C.; Konstantopoulos, G.C.; Ren, B.; Krstic, M. Improved Synchronverters with Bounded Frequency and Voltage for Smart Grid Integration. *IEEE Trans. Smart Grid* **2018**, *9*, 786–796. [CrossRef]

32.　Curi, S.; Gross, D.; Dorfler, F. Control of low-inertia power grids: A model reduction approach. In Proceedings of the 2017 IEEE 56th Annual Conference on Decision and Control (CDC), Melbourne, VIC, Australia, 12–15 December 2017. [CrossRef]

33.　Kerdphol, T.; Rahman, F.S.; Mitani, Y.; Watanabe, M.; Küfeoğlu, S. Robust virtual inertia control of an islanded microgrid considering high penetration of renewable energy. *IEEE Access* **2018**, *6*, 625–636. [CrossRef]

34.　Kerdphol, T.; Watanabe, M.; Hongesombut, K.; Mitani, Y. Self-adaptive virtual inertia control-based Fuzzy logic to improve frequency stability of microgrid with high renewable penetration. *IEEE Access* **2019**, *7*, 76071–76083. [CrossRef]

35.　Kerdphol, T.; Rahman, F.S.; Watanabe, M.; Mitani, Y. Robust virtual inertia control of a low inertia microgrid considering frequency measurement effects. *IEEE Access* **2019**, *7*, 57550–57560. [CrossRef]

36.　Ali, H.; Magdy, G.; Li, B.; Shabib, G.; Elbaset, A.A.; Xu, D.; Mitani, Y. A new frequency control strategy in an islanded microgrid using virtual inertia control-based coefficient diagram method. *IEEE Access* **2019**, *7*, 16979–16990. [CrossRef]

37.　Skiparev, V.; Belikov, J.; Petlenkov, E. Reinforcement learning based approach for virtual inertia control in microgrids with renewable energy sources. In Proceedings of the IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe), The Hague, The Netherlands, 26–28 October 2020.

38.　Rosewater, D.M.; Copp, D.A.; Nguyen, T.A.; Byrne, R.H.; Santoso, S. Battery Energy Storage Models for Optimal Control. *IEEE Access* **2019**, *7*, 178357–178391. [CrossRef]

39.　Naderipour, A.; Abdul-Malek, Z.; Hajivand, M.; Seifabad, Z.M.; Farsi, M.A.; Nowdeh, S.A.; Davoudkhani, I.F. Spotted hyena optimizer algorithm for capacitor allocation in radial distribution system with distributed generation and microgrid operation considering different load types. *Sci. Rep.* **2021**, *11*. [CrossRef]

40.　Ummels, B.C.; Gibescu, M.; Pelgrum, E.; Kling, W.L.; Brand, A.J. Impacts of Wind Power on Thermal Generation Unit Commitment and Dispatch. *IEEE Trans. Energy Convers.* **2007**, *22*, 44–51. [CrossRef]

41.　Bevrani, H. *Robust Power System Frequency Control*, 2nd ed.; Springer: Berlin, Germany, 2014.

42.　Torres, L.M.A.; Lopes, L.A.C.; Morán, T.L.A.; Espinoza, C.J.R. Self-Tuning Virtual Synchronous Machine: A Control Strategy for Energy Storage Systems to Support Dynamic Frequency Control. *IEEE Trans. Energy Convers.* **2014**, *29*, 833–840. [CrossRef]

43.　Kundur, P. *Power System Stability and Control*; MCGRAW HILL BOOK CO.: New York, NY, USA, 1994.

44.　RG-CE System Protection & Dynamics Sub Group. *Rate of Change of Frequency (RoCoF) with Stand Capability*; Technical Report; ENTSO-E: Brussels, Belgium, 2018.

45.　Kerdphol, T.; Rahman, F.; Mitani, Y.; Hongesombut, K.; Küfeoğlu, S. Virtual inertia control-based model predictive control for microgrid frequency stabilization considering high renewable energy integration. *Sustainability* **2017**, *9*, 773. [CrossRef]

46.　Sockeel, N.; Gafford, J.; Papari, B.; Mazzola, M. Virtual inertia emulator-based model predictive control for grid frequency regulation considering high penetration of inverter-based energy storage system. *IEEE Trans. Sustain. Energy* **2020**, *11*, 2932–2939. [CrossRef]

47.　Beck, H.P.; Hesse, R. Virtual synchronous machine. In Proceedings of the International Conference on Electrical Power Quality and Utilisation, Barcelona, Spain, 9–11 October 2007. [CrossRef]

48.　Guggilam, S.S.; Zhao, C.; Dall'Anese, E.; Chen, Y.C.; Dhople, S.V. Engineering inertial and primary-frequency response for distributed energy resources. In Proceedings of the 2017 IEEE 56th Annual Conference on Decision and Control (CDC), Melbourne, VIC, Australia, 12–15 December 2017. [CrossRef]

49.　Díaz-González, F.; Sumper, A.; Gomis-Bellmunt, O.; Villafáfila-Robles, R. A review of energy storage technologies for wind power applications. *Renew. Sustain. Energy Rev.* **2012**, *16*, 2154–2171. [CrossRef]

50.　Molina, M.G. *Grid Energy Storage Systems*; Wiley: New York, NY, USA, 2019; pp. 495–583. [CrossRef]

51.　Kunisch, H.J.; Kramer, K.G.; Dominik, H. Battery Energy Storage Another Option for Load-Frequency-Control and Instantaneous Reserve. *IEEE Trans. Energy Convers.* **1986**, *EC-1*, 41–46. [CrossRef]

52.　Ribeiro, P.F.; Crow, M.L. Energy storage systems for advanced power applications. *Proc. IEEE* **2001**, *89*, 1744–1756. [CrossRef]

53.　Zeng, J.; Zhang, B.; Mao, C.; Wang, Y. Use of battery energy storage system to improve the power quality and stabilty of wind farms. In Proceedings of the 2006 International Conference on Power System Technology, Chongqing, China, 22–26 October 2006. [CrossRef]

54.　Oudalov, A.; Chartouni, D.; Ohler, C. Optimizing a Battery Energy Storage System for Primary Frequency Control. *IEEE Trans. Power Syst.* **2007**, *22*, 1259–1266. [CrossRef]

55.　Bazargan, D.; Filizadeh, S.; Gole, A. Stability analysis of converter-connected battery energy storage systems in the grid. *IEEE Trans. Sustain. Energy* **2014**, *5*, 1204–1212. [CrossRef]

56.　Kerdphol, T.; Qudaih, Y.; Mitani, Y. Optimum battery energy storage system using PSO considering dynamic demand response for microgrids. *Int. J. Electr. Power Energy Syst.* **2016**, *83*, 58–66. [CrossRef]

57.　Toma, L.; Sanduleac, M.; Baltac, S.A.; Arrigo, F.; Mazza, A.; Bompard, E.; Musa, A.; Monti, A. On the virtual inertia provision by BESS in low inertia power systems. In Proceedings of the 2018 IEEE International Energy Conference (ENERGYCON), Limassol, Cyprus, 3–7 June 2018. [CrossRef]

58.　Feng, X. Dynamic balancing for low inertia power systems. In Proceedings of the 2013 IEEE Power & Energy Society General Meeting, Vancouver, BC, Canada, 21–25 July 2013. [CrossRef]

59.　Vandoorn, T.L.; Vasquez, J.C.; Kooning, J.D.; Guerrero, J.M.; Vandevelde, L. Microgrids: Hierarchical Control and an Overview of the Control and Reserve Management Strategies. *IEEE Ind. Electron. Mag.* **2013**, *7*, 42–55. [CrossRef]

60. Dragicevic, T.; Wu, D.; Shafiee, Q.; Meng, L. Distributed and decentralized control architectures for converter-interfaced microgrids. *Chin. J. Electr. Eng.* **2017**, *3*, 41–52. [CrossRef]
61. Wu, Y.; Wu, Y.; Guerrero, J.M.; Vasquez, J.C.; Li, J. AC Microgrid Small-Signal Modeling: Hierarchical Control Structure Challenges and Solutions. *IEEE Electrif. Mag.* **2019**, *7*, 81–88. [CrossRef]
62. 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*, 1935–1944. [CrossRef]
63. Corno, M.; Formentin, S.; Savaresi, S.M. Data-Driven Online Speed Optimization in Autonomous Sailboats. *IEEE Trans. Intell. Transp. Syst.* **2016**, *17*, 762–771. [CrossRef]
64. ISO. *ISO 8528-5:2005: Reciprocating Internal Combustion Engine Driven Alternating Current Generating Sets—Part 5: Generating Set*; Technical Report; International Organization for Standardization: Geneva, Switzerland, 2005.
65. ISO. *ISO 8528-5:2018 Reciprocating Internal Combustion Engine Driven Alternating Current Generating Sets—Part 5: Generating Sets*; Technical Report; International Organization for Standardization (ISO): Geneva, Switzerland, 2018.
66. ENTSO-E. *Continental Europe Significant Frequency Deviations*; Technical Report; ENTSO-E: Brussels, Belgium, 2019.
67. Wang, C.; Mi, Y.; Fu, Y.; Wang, P. Frequency Control of an Isolated Micro-Grid Using Double Sliding Mode Controllers and Disturbance Observer. *IEEE Trans. Smart Grid* **2018**, *9*, 923–930. [CrossRef]
68. Ganjian-Aboukheili, M.; Shahabi, M.; Shafiee, Q.; Guerrero, J.M. Seamless Transition of Microgrids Operation From Grid-Connected to Islanded Mode. *IEEE Trans. Smart Grid* **2020**, *11*, 2106–2114. [CrossRef]
69. Rezkalla, M.; Pertl, M.; Marinelli, M. Electric power system inertia: Requirements, challenges and solutions. *Electr. Eng.* **2018**, *100*, 2677–2693. [CrossRef]
70. Magdy, G.; Bakeer, A.; Shabib, G.; Elbaset, A.A.; Mitani, Y. Decentralized model predictive control strategy of a realistic multi power system automatic generation control. In Proceedings of the Nineteenth International Middle East Power Systems Conference, Cairo, Egypt, 19–21 December 2017. [CrossRef]
71. Kerdphol, T.; Rahman, F.S.; Watanabe, M.; Mitani, Y.; Turschner, D.; Beck, H.P. Enhanced virtual inertia control based on derivative technique to emulate simultaneous inertia and damping properties for microgrid frequency regulation. *IEEE Access* **2019**, *7*, 14422–14433. [CrossRef]
72. Tamrakar, U.; Hansen, T.M.; Tonkoski, R.; Copp, D.A. Model Predictive Frequency Control of Low Inertia Microgrids. In Proceedings of the 2019 IEEE 28th International Symposium on Industrial Electronics (ISIE), Vancouver, BC, Canada, 12–14 June 2019. [CrossRef]
73. Metzger, M.; Polakow, G. A Survey on Applications of Agent Technology in Industrial Process Control. *IEEE Trans. Ind. Inform.* **2011**, *7*, 570–581. [CrossRef]
74. Singh, V.P.; Mohanty, S.R.; Kishor, N.; Ray, P.K. Robust H-infinity load frequency control in hybrid distributed generation system. *Int. J. Electr. Power Energy Syst.* **2013**, *46*, 294–305. [CrossRef]
75. Khan, A.A.; Khan, M.Q.; Satti, S.G.; Adil, M. Robust control of hybrid distributed generation for frequency regulation. In Proceedings of the 2017 14th International Bhurban Conference on Applied Sciences and Technology (IBCAST), Islamabad, Pakistan, 10–14 January 2017. [CrossRef]
76. Fathi, A.; Shafiee, Q.; Bevrani, H. Robust Frequency Control of Microgrids Using an Extended Virtual Synchronous Generator. *IEEE Trans. Power Syst.* **2018**, *33*, 6289–6297. [CrossRef]
77. Manabe, S. Importance of coefficient diagram in polynomial method. In Proceedings of the 42nd IEEE International Conference on Decision and Control (IEEE Cat. No.03CH37475), Maui, HI, USA, 9–12 December 2003. [CrossRef]
78. Manabe, S. Coefficient diagram method in mimo application: An aerospace case study. *IFAC Proc. Vol.* **2005**, *38*, 7–12. [CrossRef]
79. Lipatov, A.V.; Sokolov, N.I. On some sufficient conditions for stability and instability of linear continuous stationary systems. *Avtomat. iTelemekh.* **1978**, *39*, 1285–1291.
80. Manabe, S. Coefficient Diagram Method as Applied to the Attitude Control of Controlled-Bias-Momentum Satellite. *IFAC Proc. Vol.* **1994**, *27*, 327–332. [CrossRef]
81. Ali, R.; Mohamed, T.H.; Qudaih, Y.S.; Mitani, Y. A new load frequency control approach in an isolated small power systems using coefficient diagram method. *Int. J. Electr. Power Energy Syst.* **2014**, *56*, 110–116. [CrossRef]
82. Coelho, J.P.; Boaventura-Cunha, J.; de Moura Oliveira, P.B. *Extended Stability Conditions for CDM Controller Design*; Springer International Publishing: Cham, Switzerland, 2015; pp. 171–182. [CrossRef]
83. Ali, H.; Magdy, G.; Xu, D. A new optimal robust controller for frequency stability of interconnected hybrid microgrids considering non-inertia sources and uncertainties. *Int. J. Electr. Power Energy Syst.* **2021**, 106651. [CrossRef]
84. Sutton, R.S.; Barto, A.G. *Reinforcement Learning: An Introduction*, 2nd ed.; The MIT Press: Cambridge, MA, USA, 2018.
85. Lillicrap, T.P.; Hunt, J.J.; Pritzel, A.; Heess, N.; Erez, T.; Tassa, Y.; Silver, D.; Wierstra, D. Continuous control with deep reinforcement learning. *arXiv* **2019**, arXiv:1509.02971.
86. Li, Y.; Gao, W.; Yan, W.; Huang, S.; Wang, R.; Gevorgian, V.; Gao, D.W. Data-Driven Optimal Control Strategy for Virtual Synchronous Generator via Deep Reinforcement Learning Approach. *J. Mod. Power Syst. Clean Energy* **2021**, 1–11. [CrossRef]
87. Guo, W.; Liu, F.; Si, J.; Mei, S. Incorporating approximate dynamic programming-based parameter tuning into PD-type virtual inertia control of DFIGs. In Proceedings of the International Joint Conference on Neural Networks, Dallas, TX, USA, 4–9 August 2013. [CrossRef]

88.   Shrestha, D.; Tamrakar, U.; Malla, N.; Ni, Z.; Tonkoski, R. Reduction of energy consumption of virtual synchronous machine using supplementary adaptive dynamic programming. In Proceedings of the 2016 IEEE International Conference on Electro Information Technology (EIT), Grand Forks, ND, USA, 19–21 May 2016. [CrossRef]

89.   Wang, D.; He, H.; Liu, D. Adaptive Critic Nonlinear Robust Control: A Survey. *IEEE Trans. Cybern.* **2017**, *47*, 3429–3451. [CrossRef] [PubMed]

90.   Mokadem, M.E.; Courtecuisse, V.; Saudemont, C.; Robyns, B.; Deuse, J. Fuzzy Logic Supervisor-Based Primary Frequency Control Experiments of a Variable-Speed Wind Generator. *IEEE Trans. Power Syst.* **2009**, *24*, 407–417. [CrossRef]

91.   Sa-ngawong, N.; Ngamroo, I. Optimal fuzzy logic-based adaptive controller equipped with DFIG wind turbine for frequency control in stand alone power system. In Proceedings of the 2013 IEEE Innovative Smart Grid Technologies-Asia (ISGT Asia), Bangalore, India, 10–13 November 2013. [CrossRef]

92.   Mentesidi, K.; Garde, R.; Aguado, M.; Rikos, E. Implementation of a fuzzy logic controller for virtual inertia emulation. In Proceedings of the 2015 International Symposium on Smart Electric Distribution Systems and Technologies (EDST), Vienna, Austria, 8–11 September 2015. [CrossRef]

93.   Hu, Y.; Wei, W.; Peng, Y.; Lei, J. Fuzzy virtual inertia control for virtual synchronous generator. In Proceedings of the 2016 35th Chinese Control Conference (CCC), Chengdu, China, 27–29 July 2016. [CrossRef]

94.   Chamorro, H.R.; Riaño, I.; Gerndt, R.; Zelinka, I.; Gonzalez-Longatt, F.; Sood, V.K. Synthetic inertia control based on fuzzy adaptive differential evolution. *Int. J. Electr. Power Energy Syst.* **2019**, *105*, 803–813. [CrossRef]

95.   Mamdani, E.; Baaklini, N. Prescriptive method for deriving control policy in a fuzzy-logic controller. *Electron. Lett.* **1975**, *11*, 625. [CrossRef]

96.   Wang, T.; Gao, H.; Qiu, J. A Combined Adaptive Neural Network and Nonlinear Model Predictive Control for Multirate Networked Industrial Process Control. *IEEE Trans. Neural Netw. Learn. Syst.* **2016**, *27*, 416–425. [CrossRef]

97.   Kennedy, J.; Eberhart, R. Particle swarm optimization. In Proceedings of the ICNN95-International Conference on Neural Networks, Perth, WA, Australia, 27 November–1 December 1995; Volume 4, pp. 1942–1948. [CrossRef]

98.   Gaing, Z.L. A Particle Swarm Optimization Approach for Optimum Design of PID Controller in AVR System. *IEEE Trans. Energy Convers.* **2004**, *19*, 384–391. [CrossRef]

99.   Magdy, G.; Bakeer, A.; Nour, M.; Petlenkov, E. A New Virtual Synchronous Generator Design Based on the SMES System for Frequency Stability of Low-Inertia Power Grids. *Energies* **2020**, *13*, 5641. [CrossRef]

100. Qi, Z.; Shi, Q.; Zhang, H. Tuning of Digital PID Controllers Using Particle Swarm Optimization Algorithm for a CAN-Based DC Motor Subject to Stochastic Delays. *IEEE Trans. Ind. Electron.* **2020**, *67*, 5637–5646. [CrossRef]

101. Frack, P.F.; Mercado, P.E.; Molina, M.G. Extending the VISMA concept to improve the frequency stability in Microgrids. In Proceedings of the 2015 18th International Conference on Intelligent System Application to Power Systems (ISAP), Porto, Portugal, 11–16 September 2015. [CrossRef]

102. Hekimoglu, B. Optimal Tuning of Fractional Order PID Controller for DC Motor Speed Control via Chaotic Atom Search Optimization Algorithm. *IEEE Access* **2019**, *7*, 38100–38114. [CrossRef]

103. Chen, J.; Yang, F.; Han, Q.L. Model-Free Predictive H Control for Grid-Connected Solar Power Generation Systems. *IEEE Trans. Control. Syst. Technol.* **2014**, *22*, 2039–2047. [CrossRef]

104. Mariethoz, S.; Fuchs, A.; Morari, M. A VSC-HVDC Decentralized Model Predictive Control Scheme for Fast Power Tracking. *IEEE Trans. Power Deliv.* **2014**, *29*, 462–471. [CrossRef]

105. Han, H.G.; Zhang, L.; Hou, Y.; Qiao, J.F. Nonlinear Model Predictive Control Based on a Self-Organizing Recurrent Neural Network. *IEEE Trans. Neural Netw. Learn. Syst.* **2016**, *27*, 402–415. [CrossRef]

106. Kou, P.; Liang, D.; Yu, L.; Gao, L. Nonlinear Model Predictive Control of Wind Farm for System Frequency Support. *IEEE Trans. Power Syst.* **2019**, *34*, 3547–3561. [CrossRef]

107. Rout, U.K.; Sahu, R.K.; Panda, S. Design and analysis of differential evolution algorithm based automatic generation control for interconnected power system. *Ain Shams Eng. J.* **2013**, *4*, 409–421. [CrossRef]

108. Sathya, M.; Ansari, M.M.T. Load frequency control using Bat inspired algorithm based dual mode gain scheduling of PI controllers for interconnected power system. *Int. J. Electr. Power Energy Syst.* **2015**, *64*, 365–374. [CrossRef]

109. Vukojevic, A.; Lukic, S. Microgrid Protection and Control Schemes for Seamless Transition to Island and Grid Synchronization. *IEEE Trans. Smart Grid* **2020**, *11*, 2845–2855. [CrossRef]

110. European Comission. *Orintations towards the First Strategic Plan for Horizon Europe*; Technical Report; European Comission: Brussels, Belgium, 2018.

111. Wang, S.; Gangammanavar, H.; Eksioglu, S.D.; Mason, S.J. Stochastic Optimization for Energy Management in Power Systems With Multiple Microgrids. *IEEE Trans. Smart Grid* **2019**, *10*, 1068–1079. [CrossRef]

112. Zeng, P.; Li, H.; He, H.; Li, S. Dynamic Energy Management of a Microgrid Using Approximate Dynamic Programming and Deep Recurrent Neural Network Learning. *IEEE Trans. Smart Grid* **2019**, *10*, 4435–4445. [CrossRef]

113. Sedhom, B.E.; El-Saadawi, M.M.; Moursi, M.E.; Hassan, M.; Eladl, A.A. IoT-based optimal demand side management and control scheme for smart microgrid. *Int. J. Electr. Power Energy Syst.* **2021**, *127*, 106674. [CrossRef]

114. Li, Y.; Liu, H.; Liu, C.; Wei, C.; Jiang, W.; Wang, F.; Wang, Z. Study on AC-side dynamic braking-based fault ride-through control for islanded renewable energy system with grid-connected VSC-HVDC transmission. In Proceedings of the 2017 Chinese Automation Congress (CAC), Jinan, China, 20–22 October 2017. [CrossRef]

115. Byrne, R.H.; Nguyen, T.A.; Copp, D.A.; Chalamala, B.R.; Gyuk, I. Energy Management and Optimization Methods for Grid Energy Storage Systems. *IEEE Access* **2017**, *6*, 13231–13260. [CrossRef]

116. Hannan, M.A.; Hussin, I.; Ker, P.J.; Hoque, M.M.; Lipu, M.S.H.; Hussain, A.; Rahman, M.S.A.; Faizal, C.W.M.; Blaabjerg, F. Advanced Control Strategies of VSC Based HVDC Transmission System: Issues and Potential Recommendations. *IEEE Access* **2018**, *6*, 78352–78369. [CrossRef]

117. Xu, Y.; Sun, H.; Gu, W.; Xu, Y.; Li, Z. Optimal Distributed Control for Secondary Frequency and Voltage Regulation in an Islanded Microgrid. *IEEE Trans. Ind. Inform.* **2019**, *15*, 225–235. [CrossRef]

118. Zidar, M.; Hatziargyriou, N.D.; Škrlec, D.; Capuder, T.; Georgilakis, P.S. Review of energy storage allocation in power distribution networks: Applications, methods and future research. *IET Gener. Transm. Distrib.* **2016**, *10*, 645–652. [CrossRef]

119. REN21. Renewables Global Futures Report: Great Debates towards 100% Renewable Energy. 2018. Available online http://www.ren21.net/future-of-renewables/global-futures-report/ (accessed on 4 March 2020).

120. Enang, W.; Bannister, C. Modelling and control of hybrid electric vehicles (A comprehensive review). *Renew. Sustain. Energy Rev.* **2017**, *74*, 1210–1239. [CrossRef]

121. Pourmousavi, S.A.; Patrick, S.N.; Nehrir, M.H. Real-Time Demand Response Through Aggregate Electric Water Heaters for Load Shifting and Balancing Wind Generation. *IEEE Trans. Smart Grid* **2014**, *5*, 769–778. [CrossRef]

122. Yang, M.; Wang, J.; An, J. Day-Ahead Optimization Scheduling for Islanded Microgrid Considering Units Frequency Regulation Characteristics and Demand Response. *IEEE Access* **2020**, *8*, 7093–7102. [CrossRef]

123. Guo, J.; Zhao, T.; Liu, W.; Zhang, J. Reliability Modeling and Assessment of Isolated Microgrid Considering Influences of Frequency Control. *IEEE Access* **2019**, *7*, 50362–50371. [CrossRef]

124. Zames, G. Feedback and optimal sensitivity: Model reference transformations, multiplicative seminorms, and approximate inverses. *IEEE Trans. Autom. Control* **1981**, *26*, 301–320. [CrossRef]

125. Yang, S.; Lei, Q.; Peng, F.Z.; Qian, Z. A Robust Control Scheme for Grid-Connected Voltage-Source Inverters. *IEEE Trans. Ind. Electron.* **2011**, *58*, 202–212. [CrossRef]

126. Mešanović, A.; Münz, U.; Heyde, C. Comparison of H∞, H2, and pole optimization for power system oscillation damping with remote renewable generation. *IFAC-PapersOnLine* **2016**, *49*, 103–108. [CrossRef]

127. Wutthithanyawat, C.; Wangnippamto, S. Design of Decentralized PID Controller with Coefficient Diagram Method Based on Inverted Decoupling for TITO System. In Proceedings of the 2018 International Electrical Engineering Congress (iEECON), Krabi, Thailand, 7–9 March 2018. [CrossRef]

128. Moharana, A.; Samarabandu, J.; Varma, R.K. Fuzzy supervised PI controller for VSC HVDC system connected to Induction Generator based wind farm. In Proceedings of the 2011 IEEE Electrical Power and Energy Conference, Winnipeg, MB, Canada, 3–5 October 2011. [CrossRef]

129. Mahmoud, M.S.; Abouheaf, M.; Sharaf, A. Reinforcement learning control approach for autonomous microgrids. *Int. J. Model. Simul.* **2019**, 1–10. [CrossRef]

130. Yi, Z.; Zhao, X.; Shi, D.; Duan, J.; Xiang, Y.; Wang, Z. Accurate Power Sharing and Synthetic Inertia Control for DC Building Microgrids With Guaranteed Performance. *IEEE Access* **2019**, *7*, 63698–63708. [CrossRef]

Appendix 4

VI

V. Skiparev, J. Belikov, E. Petlenkov, and Y. Levron. Reinforcement learning based MIMO controller for virtual inertia control in isolated microgrids. In *IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*, Novi Sad, Serbia, 2022

Reinforcement Learning based MIMO Controller for Virtual Inertia Control in Isolated Microgrids

Vjatseslav Skiparev, Juri Belikov
Department of Software Science,
Tallinn University of Technology,
Tallinn, Estonia
{vjatseslav.skiparev, juri.belikov}@taltech.ee

Eduard Petlenkov
Department of Computer Systems,
Tallinn University of Technology,
Tallinn, Estonia
eduard.petlenkov@taltech.ee

Yoash Levron
Faculty of Electrical Engineering,
Technion,
Haifa, Israel
yoashl@ee.technion.ac.il

Abstract—**In this paper, we propose a multi-input multi-output controller for optimal control of nonlinear energy storage, using deep reinforcement learning (DRL) algorithm. This controller provides the frequency support in an isolated microgrid with high penetration of variable renewable energy sources and varying system inertia. To achieve an optimal control we redesigned neural network of actor and critic, simplified deep deterministic policy gradient (DDPG) rules, and reorganized the reward/punishment system. Simulation results show the efficiency of the proposed virtual inertia control architecture in several scenarios.**

Index Terms—**Virtual inertia control, deep reinforcement learning, microgrids, renewable energy**

I. INTRODUCTION

Modern concerns about the future stability of global environment became the major vector for the development of technologies to achieve decrease in pollution [1]–[3].

The variable renewable energy (VRE) sources (e.g., wind or sun) are most available type of renewable energy. However, the integration of VRE sources at the high levels decreases the rotational inertia of the power grids and jeopardizes overall grid stability [4]–[6], which can result in disturbed balance in energy production/consumption and the frequency mismatch with rest AC power grids and local blackout(s) [1]. The energy storage system (ESS) is a technology capable to compensate the lack of rotational inertia and increase stability of power grid, which applies for energy balance in power system with renewable energy penetration [7]–[9]. The isolated microgird is a type of system, which integrates the traditional and renewable sources with various share of the power generation, and serves as an example of a power system sensitive to the VRE penetration [6]. This system has drawn significant attention and was discussed in recent works [10]–[15]. The energy balance was achieved by the energy storage, modeled using the simple first-order transfer function with output limiter, which does not provide energy accumulation. This issue was addressed in several works [16], [17], where storage was modeled as a park of electrical vehicles.

In this paper, we extend a multi-input multi-output (MIMO) deep reinforcement learning (DRL) based controller from [18] and develop an optimal control architecture for nonlinear energy storage model from [19]. We use this method for virtual inertia control of an isolated microgrid with high penetration of renewable energy sources. The proposed DRL controller is applied as an additional MIMO controller to compensate virtual inertia. Here, we propose a modified controller by redesigning actor and critic neural networks, simplifying DDPG training algorithm, and reorganizing the reward/punishment system, where the actor network is designed to provide simultaneous control of power reservation and coefficients of the positive and negative feedback loops to provide frequency support. In addition, we tested the proposed control architecture for scenarios with varying system inertia and smooth (dis)connection of renewable energy sources.

II. MOTIVATION AND PROBLEM STATEMENT

Modeling and control of isolated microgrids has several important challenges for future works. One of these challenges is the energy storage, which should simulate the natural process of energy accumulation and distribution to the grid influenced by the decreased inertia. Traditionally, ESS is used as the part of virtual inertia controller that operation principles are not optimal for inertia emulation and require integration of additional control algorithm [6], [20]. The central challenge here is to regulate the grid's frequency with the high level penetration of variable renewable energy sources. In order to provide the optimal operation of the hybrid power system it is important to develop methods for effective control of energy storage device, which provide the control of energy flow in closed systems such as isolated microgrid.

A. Structure of Studied Microgrid

In this paper, we consider model of the isolated microgrid proposed and addressed in several recent papers [10], [11], [13], [14], which incorporates power players with different inertia, see Fig. 1. The typical configuration includes residential ΔP_{RL} and industrial loads ΔP_{IL}, energy sources (mechanical power of thermal plant ΔP_m, wind turbines ΔP_W and solar power plant ΔP_{PV} combined as renewable energy ΔP_{RES}), and energy storage system. Thermal power plant
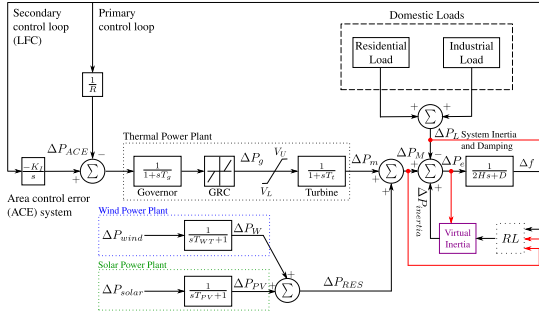
Fig. 1. Schematic representation of isloated microgrid with the proposed virtual inertia MIMO controller.

is composed of a governor with generator rate constraint ($P_{g,max}$, $P_{g,min}$) and turbine with frequency rate limiter, which restricts the valve opening/closing (V_U, V_L). The dynamic model of a microgrid utilizes the hierarchical architecture with primary and secondary control loops. The primary control loop has droop coefficient $1/R$, and the secondary loop has an area control error system ΔP_{ACE} with the second frequency controller K_I and the first-order integrator $1/s$. The frequency support is performed by a novel virtual inertia controller with output $\Delta P_{inertia}$, being the input for the total power deviation ΔP_e. Microgrid balancing system is represented as the first-order transfer function with microgrid damping coefficient D and system inertia H. The power generation by variable energy sources and loads computed as combination of two random signals with the first-order holder, their connection provides by function with smooth start. The deviation of frequency in the studied microgrid can be calculated as:

$$\Delta f = \frac{1}{2Hs + D}(\Delta P_e), \qquad (1)$$

where signals are defined as: $\Delta P_e = \Delta P_M + \Delta P_{inertia} - \Delta P_L$, $\Delta P_M = \Delta P_{RES} + \Delta P_m$, $\Delta P_m = \frac{1}{1+sT_t}\Delta P_g$, $\Delta P_g = \frac{1}{1+sT_g}(\Delta P_{ACE} - \frac{1}{R}\Delta f)$, $\Delta P_{ACE} = \frac{K_I}{s}\Delta f$, $\Delta P_L = \Delta P_{RL} + \Delta P_{IL}$, $\Delta P_{RES} = \Delta P_W + \Delta P_{PV}$. Parameters of the microgrid are summarized in Table I.

III. METHODS

A. Proposed Virtual Inertia Control Scheme

The novel virtual inertia controller is based on nonlinear control strategy of the energy storage system presented in [19], and includes the positive (red) and negative feedback (blue) loops, see Fig. 2. Positive loop has the energy accumulation part E_s with defined limits $2E_{max}$ and 0, the energy acceleration part performed as the dead-band $\zeta(E_s)$ connected to controllable signal $\Delta \alpha$, which goes to the integral block with limits λ_{max} and λ_{min}. The negative feedback loop has controllable signal $\Delta \beta$. Finally, both loops control the power output $\Delta P_{inertia}$. The control part of VIC is performed by multi-loop based controller with three output signals: two non-negative $\Delta \alpha$ and $\Delta \beta$, and one real $\Delta \gamma$, which controls the

TABLE I
NOMENCLATURE: PARAMETERS OF A MICROGRID

Parameter	Physical meaning	Nominal value	Unit
T_t	time constant of the turbine	0.4	s
T_g	time constant of the governor	0.1	s
K_I	integral control variable gain	0.05	s
H	system inertia	0.083	p.u.MW s
D	system damping coefficient	0.015	p.u.MW s
R	droop characteristic	2.4	Hz/p.u.MW
T_{WT}	time constant of wind turbines	1.85	s
T_{PV}	time constant of solar system	1.5	s
V_U	maximum limit of valve gate speed	0.5	–
V_L	minimum limit of valve gate speed	0	–
$P_{g,max}$	maximum generation rate	0.12	p.u. MW/min
$P_{g,min}$	minimum generation rate	-0.12	p.u. MW/min
$\zeta(E_s)_{max}$	maximum ESS rate	1	–
$\zeta(E_s)_{min}$	minimum ESS rate	-1	–
E_{max}	maximum capacity of ESS	1.0	–
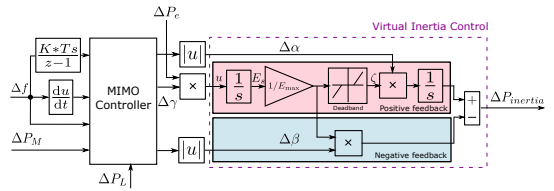
Fig. 2. Schematic representation of the proposed virtual inertia control strategy.

power inflow ΔP_e from the microgrid to the storage device. The proposed multi-loop DRL controller has several inputs Δf, ΔP_M and ΔP_L. The Δf value depends on system inertia and its stability is important. Therefore, DRL controller includes the Δf acceleration via differential block and the total deviation by discrete-time integration block. Considered microgrid model has the energy storage model (see Fig. 2), which provides the power accumulation and (dis)charge dynamics. Dynamics of energy storage is defined as:

$$\frac{dE_s}{dt} = u(t) = \Delta P_e \Delta \gamma,$$
$$E_s(0) = 0.5E_{max}, \ 0 \le E_s \le E_{max}, \qquad (2)$$

where ΔP_e is the power flowing into the storage and $\Delta \gamma$ is the control law. The dead-zone function is defined as:

$$\zeta(E_s) = \begin{cases} E_{max}, & \frac{E_s}{E_{max}} > \zeta(E_s)_{max} \\ 0, & \zeta(E_s)_{max} \ge \frac{E_s}{E_{max}} \ge \zeta(E_s)_{min} \\ -E_{max}, & \frac{E_s}{E_{max}} < \zeta(E_s)_{min} \end{cases} \qquad (3)$$

where the input is defined as $\frac{d\lambda}{dt} = \Delta \alpha \zeta(E_s)$, $\lambda(T) = 0$. The following part represents ESS controllable power output that summarizes all signals:

$$\Delta P_{inertia} = \frac{d\lambda}{dt} - \Delta \beta \frac{E_s}{E_{max}}. \qquad (4)$$

Fig. 3. Schematic representation of implemented deep reinforcement learning control algorithm.

B. Deep Reinforcement Learning based Controller

Proposed deep reinforcement learning based controller is a combination of two artificial neural networks: the actor $\mu(s \mid \theta_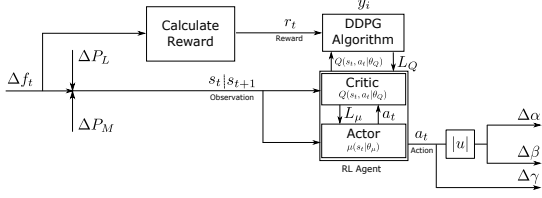\mu)$ that preforms the frequency control and the critic $Q(s, a \mid \theta_Q)$ that selects the best actions of an actor (see Fig. 3). The training rules of agent are defined by certain reinforcement learning (RL) policy (e.g. Dynamic programming, Monte Carlo, Q-learning, SARSA, DDPG, A3C, SAC) [21]–[23]. The major feature of RL training is a strategy based on the rule of try and error. When the actor makes a right action a_t according to the designed reward/punishment rules, it receives a reward r at each step of the action, otherwise take a punishment $-r$. The training duration depends on the number of the epochs and learning rate α_μ actor and α_Q critic networks, after each epoch the rewarding system calculates average reward r_{avr}. If the average reward of last epoch is significantly less than previous, then RL repeat weights initialization of neural networks to try new control strategy, unless the accumulated reward will reach a target value r_{target}. We apply this architecture as additional MIMO controller of VIC scheme, which is illustrated in Fig. 2.

C. Simplified Deep Deterministic Policy Gradient

We briefly explain the key application details of the proposed controller based on RL algorithm optimized by simplified deep deterministic policy gradient (SDDPG). Original deep deterministic policy gradient is model-free reinforcement learning algorithm, designed for tasks with low-dimensional continuous action space [22]. The DDPG optimization combines Deep Q-learning (DQL) [23] and deterministic policy gradient (DPG) [21], [22], which inherits from the DQL neural actor-critic architecture. In contrast to original DDPG, the modified version has simplified loss function for actor L_μ and training is performed without replay buffer. The principle of optimization is based on the search of a minimal difference between target action-value function y_i and critic network reward prediction $Q(s_i, a_i \mid \theta_Q)$ for each actor network $\mu(s \mid \theta_\mu)$ decision a_t (i.e., $\Delta\alpha$, $\Delta\beta$, $\Delta\gamma$) per step i in microgrid state s_t (i.e., Δf_t, ΔP_L, ΔP_M). The aim of simplified DDPG algorithm is to receive the maximum possible average reward r_t per training episode via minimization of critic L_Q and actor L_μ loss functions. Pseudo-code shown in Algorithm 1 illustrates simplified DDPG algorithm.

Algorithm 1 Simplified DDPG algorithm
___
1: Initialize critic $Q(s, a \mid \theta_Q)$ and actor $\mu(s \mid \theta_\mu)$ networks with random weights θ_Q and θ_μ.
2: Initialize learning rate for critic $\alpha_Q = 0.1$ and actor $\alpha_\mu = 0.1$ networks.
3: Initialize smooth $\tau = 0.5$ and discount factor $\gamma = 0.25$.
4: **for** $episode = 1$ to M **do**
5:     **for** $t = 1$ to T **do**
6:         Receive initial process observation as state s_t.
7:         Select action $a_t = \mu(s_t \mid \theta_\mu)$ according to current reward prediction $Q(s_i, a_i \mid \theta_Q)$.
8:         Execute action a_t, observe reward r_t and future state s_{t+1}.
9:         Select the best critic reward prediction weights $Q'(s_{i+1}, \mu'(s_{i+1}|\theta_{\mu'})$ according to explored max average reward $\max(\frac{1}{N}\sum_i r_i(s_i, a_i))$
10:        Set action-value function of DDPG policy $y_i = r_i + \gamma Q'(s_{i+1}, \mu'(s_{i+1}|\theta_{\mu'}) \mid \theta_{Q'})$.
11:        Update critic by minimizing the loss: $L_Q = \frac{1}{N}\sum_i (y_i - Q(s_i, a_i \mid \theta_Q))^2$.
12:        Update actor policy using sampled gradient of critic: $L_\mu = \frac{1}{N}\sum_i (Q(s, a \mid \theta_Q)|_{s=s_i, \alpha=\mu(s_i)})^2$
13:        Update target networks:
14: $\theta^{Q'} \leftarrow \tau\theta^Q + (1-\tau)\theta^{Q'}\alpha_Q L_Q$,
15: $\theta^{\mu'} \leftarrow \tau\theta^\mu + (1-\tau)\theta^{\mu'}\alpha_\mu L_\mu$.
16:     **end for**
17: **end for**
___

D. RL Agent Reward/Punishment System

The reward/punishment system proposed in [18] was redesigned to provide a more accurate punishment system for DRL agent in virtual inertia control task. Reward system is organized to transform the measured frequency deviation and other signals to the reward/punishment r_t, defined as:

$$r_t = \begin{cases} \frac{1}{0.5+|\Delta f|}, & \text{for } \Delta f < 0.05, \\ -2|\Delta f|, & \text{for } \Delta f > 0.05, \\ 0.1\alpha, & \text{for } \alpha > \beta, \\ -0.5\beta, & \text{for } \beta > \alpha. \end{cases} \quad (5)$$

To provide instructions for Δf regulation system is organized as follows: initial signal transforms to the absolute value $|\Delta f|$, if $\Delta f < 0.05$ the DRL agent receive a reward, otherwise system does a punishment. In order to provide reasonable rewarding of each action, the reward is limited by $\Delta f \in [0.05, 2]$; however, the punishment is unlimited and multiplied by constant value 2. Additional reward rules are provided to help RL agent understand the difference between $\Delta\alpha$ and $\Delta\beta$ magnitudes.

IV. NUMERICAL RESULTS

The renewable energy sources are modeled using $\Delta P_{wind} \in [0.1, 0.125]$ and $\Delta P_{solar} \in [0.075, 0.125]$. Load is modeled using $\Delta P_{RL} \in [0.15, 0.25]$ and $\Delta P_{IL} \in [0.2, 0.325]$. We organize experiments with and without MIMO controller to

show how proposed methodology increases the ESS operation efficiency. In addition, in each scenario we perform experiments with different system inertia 100% (i.e. $H = 0.083$) and 40% (i.e. $H = 0.032$) to show how it affects on energy storage charge dynamics. Simulation results of the proposed virtual inertia controller are illustrated in Figs. 4-7.

Scenario 1: The first simulation illustrates the nominal case without any changes during the operation time, see Fig. 4. The proposed virtual inertia controller slightly influences the E_s, and it can be seen how different levels of inertia affect the process of energy accumulation. According to the obtained results, lower inertia causes the lowered energy accumulation in ESS. If the energy storage is almost discharged, then frequency support of microgrid becomes more difficult, but still frequency deviation is kept in the range ± 0.05 Hz. In contrast, without additional controller the deviation is higher. Figure 5 illustrates performance of the proposed controller with three controllable signals.

Fig. 5. Control signals for Scenario 1.
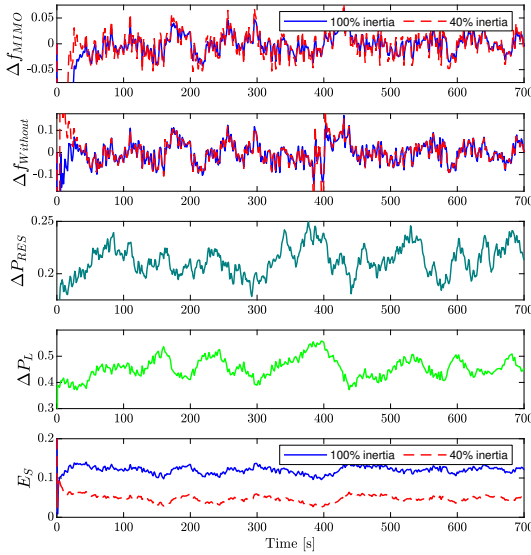
Fig. 4. Simulation results for Scenario 1.

Scenario 2: The second scenario provides the smooth connection of wind turbines at $t = 300$ s and disconnection at $t = 400$ s. It can be seen that ESS charges faster after connection of the wind power and starts discharge after disconnection of this power source as depicted in Fig. 6. As we can see before connection of wind turbines energy storage has a minimal state of charge. After disconnection of wind turbines the energy storage provides active frequency support of microgrid, helping to stabilize Δf in case of the proposed controller; however, in case without additional controller frequency keeps oscillating. In contrast to previous scenario the influence of low inertia on energy reservation is
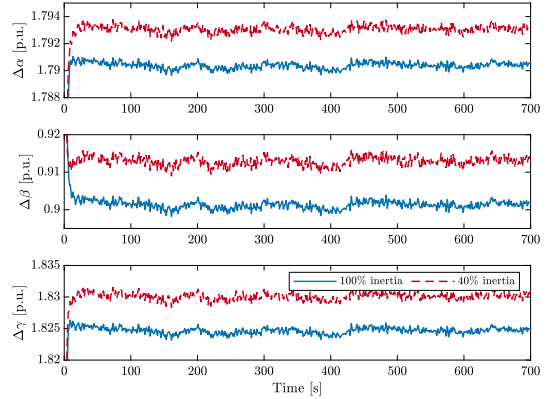
higher. Figure 7 shows how controller responds for connection of renewable source with excess power income.
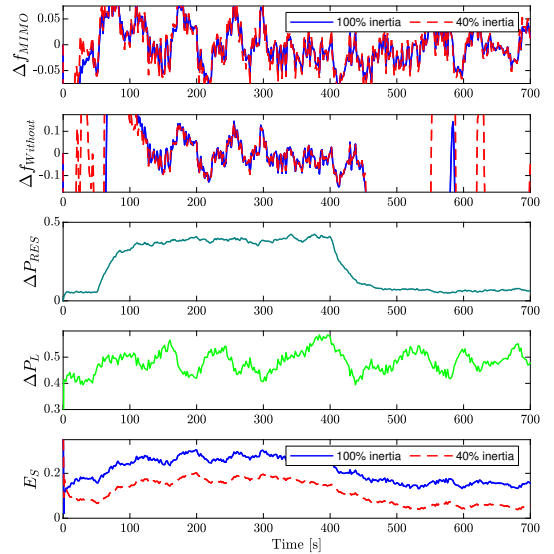
Fig. 6. Simulation results for Scenario 2.

Tables II and III show evaluation results of the proposed controller in nominal scenario and scenario with (dis)connecting renewable energy sources. Root mean square error (RMSE) and integral absolute error (IAE) statistical metrics are used to validate the performance. Both metrics confirm that the proposed VIC and DRL based controller performs better than the standard one.
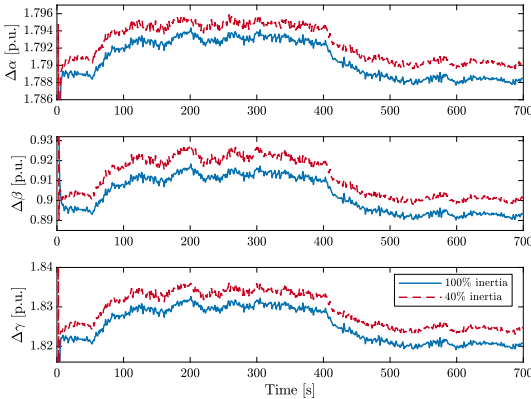
Fig. 7. Control signals for Scenario 2.

TABLE II
PERFORMANCE OF THE PROPOSED CONTROLLER: IAE

Approach	With MIMO controller		Without MIMO controller	
Inertia	100%	40%	100%	40%
Scenario 1	23.5960	24.7705	44.1569	51.1119
Scenario 2	37.4870	38.7273	2040.9	1725.6

TABLE III
PERFORMANCE OF THE PROPOSED CONTROLLER: RMSE

Approach	With MIMO controller		Without MIMO controller	
Inertia	100%	40%	100%	40%
Scenario 1	0.0837	0.0842	0.1446	0.2349
Scenario 2	0.0999	0.1027	3.0616	2.9267

V. DISCUSSION AND CONCLUSIONS

In this paper, we propose a novel virtual inertia control scheme that manages the power flow in energy storage system within an isolated microgrid.

We applied the modified deep reinforcement learning based controller as optimal MIMO controller for the proposed VIC. In order to receive a robust control architecture, the structure of neural actor and critic were changed. In addition, we simplified the policy rules of the DDPG algorithm and redesigned the reward/punishment system. To validate the proposed algorithm we performed simulations with the nominal scenario and with smooth (dis)connection of renewable energy sources for the cases of nominal (100%) and decreased (40%) inertia in each scenario. It was shown that the proposed controller is capable of providing good response and frequency support in all simulated scenarios.

REFERENCES

[1] B. Kroposki, B. Johnson, Y. Zhang, V. Gevorgian, P. Denholm, B.-M. Hodge, and B. Hannegan, "Achieving a 100% renewable grid: Operating electric power systems with extremely high levels of variable renewable energy," *IEEE Power and Energy Magazine*, vol. 15, no. 2, pp. 61–73, 2017.

[2] F. Lamnabhi-Lagarrigue, A. Annaswamy, S. Engell, A. Isaksson, P. Khargonekar, R. M. Murray, H. Nijmeijer, T. Samad, D. Tilbury, and P. V. den Hof, "Systems & control for the future of humanity, research agenda: Current and future roles, impact and grand challenges," *Annual Reviews in Control*, vol. 43, pp. 1–64, 2017.

[3] F. Milano, F. Dörfler, G. Hug, D. J. Hill, and G. Verbic, "Foundations and challenges of low-inertia systems (invited paper)," in *2018 Power Systems Computation Conference (PSCC)*, 2018.

[4] M. Dreidy, H. Mokhlis, and S. Mekhilef, "Inertia response and frequency control techniques for renewable energy sources: A review," *Renewable and Sustainable Energy Reviews*, vol. 69, no. 1, pp. 144–155, 2017.

[5] H. T. Nguyen, G. Yang, A. H. Nielsen, and P. H. Jensen, "Combination of synchronous condenser and synthetic inertia for frequency stability enhancement in low-inertia systems," *IEEE Transactions on Sustainable Energy*, vol. 10, no. 3, pp. 997–1005, 2019.

[6] V. Skiparev, R. Machlev, N. R. Chowdhury, Y. Levron, E. Petlenkov, and J. Belikov, "Virtual inertia control methods in islanded microgrids," *Energies*, vol. 14, no. 6, p. 1562, 2021.

[7] A. Ulbig, T. S. Borsche, and G. Andersson, "Impact of low rotational inertia on power system stability and operation," in *The International Federation of Automatic Control IFAC Cape Town*, 2014.

[8] A. Serpi, M. Porru, and A. Damiano, "An optimal power and energy management by hybrid energy storage systems in microgrids," *Energies*, vol. 10, no. 11, p. 1909, 2017.

[9] D. M. Rosewater, D. A. Copp, T. A. Nguyen, R. H. Byrne, and S. Santoso, "Battery energy storage models for optimal control," *IEEE Access*, vol. 7, pp. 178 357–178 391, 2019.

[10] H. Ali, G. Magdy, B. Li, G. Shabib, A. A. Elbaset, D. Xu, and Y. Mitani, "A new frequency control strategy in an islanded microgrid using virtual inertia control-based coefficient diagram method," *IEEE Access*, vol. 7, pp. 16 979–16 990, 2019.

[11] G. Magdy, G. Shabib, A. A. Elbaset, and Y. Mitani, "A novel coordination scheme of virtual inertia control and digital protection for microgrid dynamic security considering high renewable energy penetration," *IET Renewable Power Generation*, vol. 13, no. 3, pp. 462–474, 2019.

[12] T. Kerdphol, F. S. Rahman, M. Watanabe, and Y. Mitani, "Robust virtual inertia control of a low inertia microgrid considering frequency measurement effects," *IEEE Access*, vol. 7, pp. 57 550–57 560, 2019.

[13] V. Skiparev, J. Belikov, and E. Petlenkov, "Reinforcement learning based approach for virtual inertia control in microgrids with renewable energy sources," in *2020 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe)*, The Hague, NL, 2020.

[14] N. Sockeel, J. Gafford, B. Papari, and M. Mazzola, "Virtual inertia emulator-based model predictive control for grid frequency regulation considering high penetration of inverter-based energy storage system," *IEEE Transactions on Sustainable Energy*, vol. 11, no. 4, pp. 2932–2939, 2020.

[15] T. Kerdphol, F. S. Rahman, Y. Mitani, M. Watanabe, and S. Kufeoglu, "Robust virtual inertia control of an islanded microgrid considering high penetration of renewable energy," *IEEE Access*, vol. 6, pp. 625–636, 2018.

[16] M.-H. Khooban, "Secondary load frequency control of time-delay stand-alone microgrids with electric vehicles," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 9, pp. 7416–7422, 2018.

[17] G. Magdy, H. Ali, and D. Xu, "A new synthetic inertia system based on electric vehicles to support the frequency stability of low-inertia modern power grids," *Journal of Cleaner Production*, vol. 297, p. 126595, 2021.

[18] V. Skiparev, J. Belikov, and E. Petlenkov, "Mimo reinforcement learning based approach for frequency support in microgrids with high renewable energy penetration," in *IEEE PES General Meeting*, 2021.

[19] Y. Levron and J. Belikov, "Control of energy storage devices under uncertainty using nonlinear feedback systems," in *IEEE PES General Meeting*, 2020.

[20] T. Kerdphol, F. S. Rahman, M. Watanabe, and Y. Mitani, *Virtual Inertia Synthesis and Control*. Springer International Publishing, 2021.

[21] D. Silver, G. Lever, N. Heess, T. Degris, D. Wierstra, and M. Riedmiller, "Deterministic policy gradient algorithms," in *31st International Conference on Machine Learning*, vol. 22, Beijing, China, 2014.

[22] 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," 2016, [Online]. Available: arXiv:1509.02971v6.

[23] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, 2nd ed. The MIT Press, 2018.

Appendix 5

V
K. Nosrati, V. Skiparev, A. Tepljakov, E. Petlenkov, Y. Levron, and J. Belikov. Coordinated PI-based frequency deviation control of isolated hybrid microgrid: An online multi-agent tuning approach via reinforcement learning. In *2022 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*, pages 1–5, 2022

Coordinated PI-based frequency deviation control of isolated hybrid microgrid: An online multi-agent tuning approach via reinforcement learning

K. Nosrati, A. Tepljakov, E. Petlenkov
Department of Computer Systems
Tallinn University of Technology
Tallinn, Estonia
{komeil.nosrati, aleksei.tepljakov,
eduard.petlenkov}@taltech.ee

Y. Levron
Faculty of Electrical Engineering
Technion
Haifa, Israel
yoashl@ee.technion.ac.il

V. Skiparev, J. Belikov
Department of Software Science
Tallinn University of Technology
Tallinn, Estonia
{vjatseslav.skiparev, juri.belikov}@taltech.ee

Abstract—Numerous remote area applications welcome stand-alone renewable energy power generation systems or isolated microgrids (MGs). Due to the nature of solar and wind energy, the frequency deviation control (FDC) in hybrid MGs has become more complicated and critical than the conventional grid for power quality purposes. By using a coordination control strategy between a double-layered capacitor and a fuel cell, our mission here is to design a FDC system based on the PI controller which is tuned by an artificial neural network (ANN) in a multi-agent structure. To achieve this aim, a reinforcement learning technique is applied to train the ANN-based tuners. The performance of the proposed FDC system has been verified under different conditions by using real data to demonstrate the stability and robustness of the proposed controller.

Index Terms—Microgrid, frequency deviation control, multi-agent, neural networks, reinforcement learning

I. INTRODUCTION

Wind turbine generator (WTG) and photovoltaic (PV) systems are two of the most promising renewable technologies in the hybrid stand-alone power generation systems and isolated microgrids (MG) utilized in many remote area applications with no access to grid electricity [1]. Energy storage systems (ESSs) and additional power generation components such as batteries and fuel cells (FCs) are integrated into the isolated MG to mitigate power fluctuation caused by changing weather conditions and to provide a stable power supply [2].

FC system, a high efficient and modular energy generation element using chemical energy conversion, is one of the most reliable energy technologies for sustainable future [3]. Parallel to its good power capability in steady-state operation, the main drawback of this system is its slow dynamics and time delay [4]. Therefore, it is always utilized along with ESSs to change its power to a desired value and improve system performance. Due to the restricted control operation

of hydrogen in FC and its long time delay, the usual DC link capacitor cannot compensate for the variation of load demand [5]. So, double-layered capacitors (DLCs) with fast power response and associated with DC/DC converters can complement the slower power output of the main source to compensate for the load demand variation and transient response of the FC system [6].

Unlike a single power system, there are two or more power generation sources in a hybrid MG (HMG) that enhance the operating characteristics of the system [7]. A proper coordination of frequency deviation control (FDC) of components can ensure effective power delivery to the load sides. Among different efforts on FDC of HMG, the small signal stability analysis of a hybrid power system with an isolated load was discussed in [8]. In [9] and by using reduced dump load technique, the load FDC of an isolated power plant was achieved. As a contribution to FDC and power quality enhancement, a coordinated proportional-integral-differential (PID) was given in an isolated HMG [10], in which some improvements were proposed in terms of time domain response and MSE value as compared with previous works.

By optimal parameter setting of the PID controller family, an efficient stability condition of grid operation can be achieved. In this regard, some control and optimization approaches, such as genetic algorithm (GA) [11], fractional order control [12], biogeography-based optimization (BBO) technique [13], and multi-verse optimization (MVO) technique [14] were proposed for the FDC of HMG. However, all previous studies attempted to achieve a higher degree of FDC by adjusting parameters in an offline approach, associated with a few problems, including premature convergence, time complexity, and parameter tuning search space. The control options in these works are limited to manual tuning, and all of the controller's features, such as the use of integration and derivative actions, cannot be manipulated automatically.

In this study, we aim to provide the synthesis of a new flexible PI-based FDC to evaluate the grid stability under different levels of renewable energy systems (RESs) penetration

and load disturbances. For this purpose, we propose an online tuning of the PI controller based on an artificial neural network (ANN), which is trained by the reinforcement learning (RL) algorithm. This automatic tuning technique, which admits to capturing all the remarkable advantages of the PI controller, including robustness improvement and disturbance rejection, can give rise to the generation of control laws that lead to superior performance of the system. The research motivation here addresses the problem of FDC for an islanded HMG based on an automatic tuned ANN-PI controller associated with stochastic RL (SRL) and pursues the development of an online multi-agent strategy to improve the system performance. It is worth mentioning that the proposed algorithm shows a successful tuning effort of the controller in a multi-agent structure, which is valid for FDC purposes in HMG typologies.

II. System Configuration and Modeling

A. System Structure

The proposed isolated power generation/energy storage system, or HMG, is depicted in Fig. 1. The power generation subsystems comprise a wind turbine, PV panels, a FC system, and a DLC bank employed as ESS. All four systems are connected in parallel to a common AC bus line to supply isolated loads with PV, FC, and DLC connected via three separate AC/DC converters. In the proposed system, RESs PV and WTG are used as primary energy generators, with priority given to power generation to meet load demand.

To study the proposed HMG system, high order mathematical models with nonlinear dynamics must be used for each subsystem. In our scenario, simplified models such as linear first-order transfer functions are often used for simulating and investigating all components of such MGs. As a result, the system non-linearities will be ignored, and the system simulations in this study will use a simplified model of the components. In what follows, mathematical models of the components are presented.

B. WTG Model

Wind speed has a direct effect on the output power of WTGs, as mechanical power in the turbine section is determined by

$$P_{WTG} = \frac{1}{2}\rho A v^3 C_p(\lambda, \theta), \tag{1}$$

where ρ is the density of air, A is the area swept by the blades, v is the wind velocity, C_P is the rotor efficiency, which is a function of tip speed ratio λ and pitch angle θ.

As depicted in Fig. 2a, the $C_{P,max}$ can be acquired for a given direction of the blades and when λ is in its special value, which directly depends on the turbine aerodynamic structure. Depending on the wind speed, the rotor speed can keep λ at its optimal level, which means that the most energy from the wind can be used. Also, Fig. 2b illustrates the variation of the output power according to the wind speed. The power remained constant by using the pitch angle control system to prevent excessive rotor speed and preserve the equipment
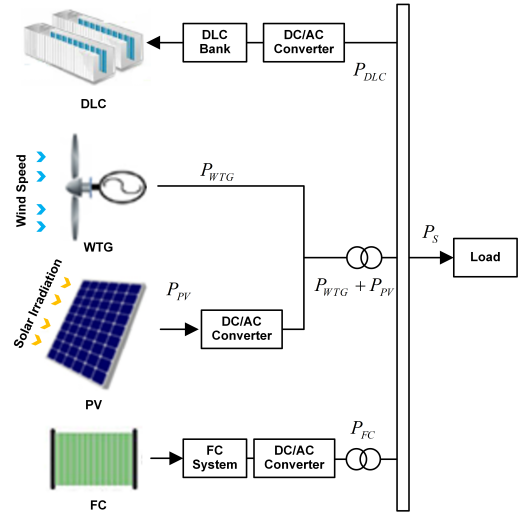
Fig. 1. General configuration of the isolated HMG.

when v increases through the rated wind velocity. When v is smaller and greater than the cut out and on wind speeds, respectively, i.e., $14 < v < 25$ m/s, the output power takes its constant maximum value, and is zero for $v < 4$ m/s. However, for $v > 25$ m/s, the system takes out of operation.
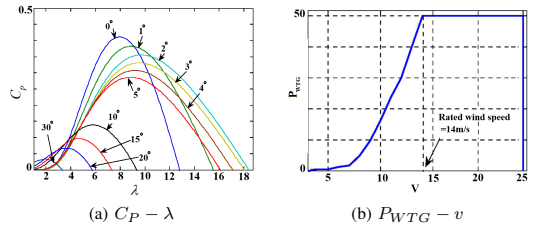
Fig. 2. Variation of C_P and P_{WTG} according to the λ and v.

By defining T_{WTG} as a time constant, the dynamical model of WTG in the frequency domain can be given by

$$\frac{\Delta P_{WTG}}{\Delta P_{wind}} = \frac{1}{sT_{WTG} + 1}, \tag{2}$$

where ΔP_{wind} and ΔP_{WTG} are the variations of the mechanical power and output power of WTG, respectively.

C. PV Model

One of the most promising, flexible, and environmentally friendly power sources is the PV system, which consists of PV panels connected in series and parallel structures. This system converts solar radiation to electrical data based on solar cell temperature and array area, and its output power P_{PV} can be

expressed as follows in terms of conversion efficiency η, PV array area S, solar irradiation φ, and ambient temperature T_a:

$$P_{PV} = \eta S \varphi (1 - 0.005(T_a + 25)). \tag{3}$$

The variations of solar irradiation to output power of PV system can be described as

$$\frac{\Delta P_{PV}}{\Delta \varphi} = \frac{1}{sT_{PV} + 1}, \tag{4}$$

where T_a is the PV time constant.

D. FC Model

The FC technology, which has been considered as a high-efficiency power generation device, can convert chemical energy into electrical energy by using hydrogen and oxygen. This static device has a slow dynamics in its fuel supply sections such as pumps and valves, resulting in slower power output. The dynamic response of the FC system can be given by introducing the first-order time delay transfer function and T_{FC} as the FC time constant

$$G_{FC} = \frac{1}{sT_{FC} + 1}. \tag{5}$$

E. DLC Model

General Electric Company first patented electro-chemical capacitors in 1957, consisting of porous electrodes using the DLC mechanism of charging [15]. Nowadays, this technology plays a key part in fulfilling the demands of possible applications, including memory back-up, electric vehicles, power quality, and RESs. The transfer function of this highly efficient device with a fast load frequency can be represented by the first-order lag equation

$$G_{DLC} = \frac{1}{sT_{DLC} + 1}, \tag{6}$$

where T_{DLC} is a time constant.

F. Frequency Deviation System

For the stable operation of the HMG system with different power generations, effective control of components is necessary. This can be achieved by regulating the fluctuation in the frequency profile Δf, which can be expressed using the power balance $\Delta P = P_{Net} - P_{Load}$ as

$$\Delta f = \frac{\Delta P}{K_{HMG}}, \tag{7}$$

where P_{Net} and P_{Load} are the net and load power, respectively, and K_{HMG} is the system frequency constant of HMG. The dynamical model of frequency variation according to per unit power deviation can be given by

$$\frac{\Delta f}{\Delta P} = \frac{1}{sM + D}, \tag{8}$$

where M and D are the inertia and damping constants of HMG, respectively, see [15].

III. Proposed Multi-agent Control Strategy

In the given HMG system, the RESs, including PV and WTG, are considered the main and primary energy sources to generate power for the loads. Because of the high dependency of the produced power by these two subsystems on weather parameters, the HMG is equipped by FC to provide the required energy to meet the connected loads. To obviate the slow dynamical response of the FC, the residual energy of the HMG is provided by ESSs, especially the DLC system, which is flexible and fast. This subsystem can effectively complement the slower power output of the main source to compensate for the variation of load demand and FC system power.

By using a coordinated control strategy between FC and DLC (see Fig. 3), the FDC problem can be solved with enhanced power quality. In this regard, DLC and FC systems act as backup systems and compensate for high and low frequency deviation, respectively. A high-pass filter (HPF) here can reduce the charging and discharging of DLC in long-term operation. Also, as seen from Fig. 3, the net power generation is comprised by the following equation

$$P_{Net} = P_{WTG} + P_{PV} + P_{FC} \pm P_{DLC}. \tag{9}$$

Here, an online and automatic tuning based PI controller, which admits to capturing all the remarkable advantages of the controller, including robustness improvement and disturbance rejection, is applied for the control purpose. The robustness of the isolated HMG system can be improved using an online tuning of the PI controller based on a NN technique trained by the SRL algorithm, resulting in superior system performance. It shows that the controller can be tuned in a multi-agent structure, which is good for FDC in HMG types of systems. This online-tuned ANN-PI controller strategy shows a successful effort to do so as shown in the next sections.

A. ANN Tuning-based PI Controller

In the proposed solution, each ANN performs online tuning of each PI controller simultaneously. After each HMG disturbance, we provide a reasonable change in the controller's two parameters tuning, whose transfer function is stated as $K_{PI} = K_p(t) + K_i s^{-1}(t)$. In the proposed architecture, we avoid tuning of all PID controller coefficients because the differential part is sensitive to disturbances and difficult to change without losing tone stability. To organize effective and fast training of two independent agents, we designed the multi-agent SRL optimization method, where weight range is designed for each agent individually. We propose a simple ANN with positive weights in the range of $[0, 5]$ and $[0, 10]$ for agents 1 and 2, respectively, and a modified `tansig` activation function by adding absolute value

$$y = \frac{m}{1 + e^{-|0.1x|}} - n, \tag{10}$$

where the absolute value of x is used to avoid negative output, provide a sufficient range of coefficients for the PI controller, and make the tuning more robust. For effective coordination purposes of the two subsystems, DLC and FC, the parameters

m and n are considered to be 200 and 100 for the DLC element, and 10 times larger for the FC subsystem. We avoid connecting a high-pass filter (HPF) to the ANN-based tuner since it has a negative impact on the stability of the tuner and confuses the reward system. Therefore, the input signal for both agents is the same.
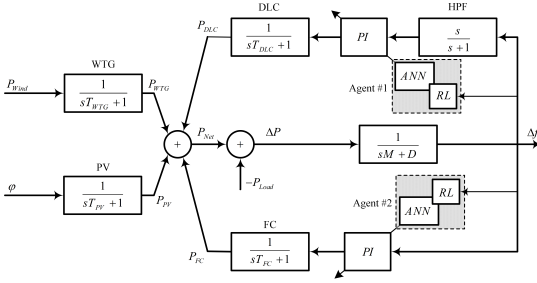

Fig. 3. Proposed coordinated multi-agent FDC scheme.

B. SRL-based Training for Multiple Agents

We use a basic SRL to train multiple RL agents and provide optimal online tuning of each PI controller, in which random weights are generated M times. The training is based on reward feedback organized by reward system proposed in [16], where the best weights can be obtained via the selection of actions with the highest possible reward. In this approach, we perform actions a_t (i.e., $K_p(t)$ and $K_i(t)$) and check the received maximum average reward $r_{avr,max}$ of each agent individually, where the main criterion for rewarding is the magnitude of Δf (see (11)), considered as control error, and an agent receives a positive reward $+r$ if frequency deviation is less than band 0.01 Hz; otherwise, it takes a negative reward or punishment $-r$. The major problem with multiple agents is synchronized training, because the actions of the first agent affect the results of the second agent and *vice-versa*. Therefore, we admit that the reward of both agents is easier to calculate by a common summarized reward and saving weights with the best reward for each agent individually. The following equation illustrates the provided common reward rules for all agents:

$$r_t = \begin{cases} \frac{1}{0.5+10|\Delta f|}, & \text{if } 10|\Delta f| < 0.01 \\ -20|\Delta f|, & \text{if } 10|\Delta f| > 0.01. \end{cases} \quad (11)$$

IV. NUMERICAL RESULTS

To develop the proposed multi-agent management strategy and to investigate the coordinated ANN-PI controller performance, the given control scheme (see Fig. 3) has been implemented using MATLAB/Simulink with model parameters $T_{WTG} = 1.5$ s, $T_{PV} = 1.8$ s, $T_{FC} = 0.26$ s, $T_{DLC} = 0.01$ s, $M = 0.4$, and $D = 0.03$. Figure 4 (top and middle plots) shows the output power of WTG and PV systems, respectively; while the bottom plot depicts the load demand applied for

---

Algorithm 1 Multi-Agent SRL optimization approach.

1: Initialize neural network of agent n as $\mu_n(s \mid \theta_{\mu_n})$
2: **for** $episode = 1$ to M **do**
3:     Initialize random weights of each agent θ_{μ_n}
4:     **for** $i = 1$ to N **do**
5:         Execute action $a_{t_n} = K_{PI_n}(t)$, observe reward r_{t_n} at each step i in common state $s_t = \Delta f_t$
6:         **if** Average reward of n agent $r_{avr_n} > r_{avr,\max_n} = \max\left(\frac{1}{N}\sum_i r_{i_n}(s_{i_n}, a_{i_n})\right)$ **then**
7:             Save weights of each agent θ_{μ_n}
8:         **end if**
9:     **end for**
10: **end for**

HMG systems. Figure 5 shows how the FDC problem can be appropriately solved by using the proposed controller, in which whole hybrid power generation is better compensated by considering the effects of system frequency variation as compared to the usual PI controllers and when the system is free of control (without control). In terms of transient response, ANN-PI has a smaller overshoot with a shorter settling time, indicating a faster transient time. In terms of steady state, the proposed approach has small steady-state error and reduces the values of the integral absolute error (IAE), mean square error (MSE), and root MSE (RMSE) when compared to the traditional PI controller for the entire period (see Table I).

Fig. 4. Output powers of WTG, PV, and load variations.
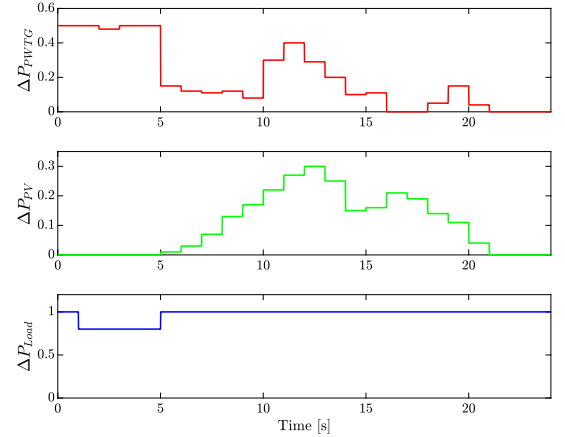
TABLE I
PERFORMANCE RESULTS.

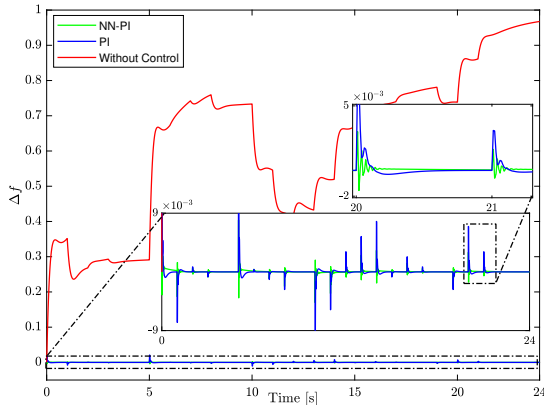| Controller	IAE	RMSE	MSE
Coordinated ANN-PI	0.0027	0.0004	0.0000016
Coordinated PI	0.005	0.0011	0.000012
No control	14.93	0.6595	0.435

Fig. 5. FDC results of the proposed control approach (ANN-PI) compared to the classical control (PI), and when HMG is free of control (without control).

According to the results for the proposed controller, we can see an insignificant influence of WTG and PV disturbances on the proposed combination of ANN and PI controller. In contrast, the traditional PI controllers show less flexibility in the distributive process. Figure 6 illustrates the parameters of the online tuned PI controllers. Here, we can see how ANN adjusts the parameters of both controllers to optimal values after each step in the signal. For the usual PI controllers, the parameters are fixed at $K_{p,DLC} = 30$, $K_{i,DLC} = 0.1$, and $K_{p,FC} = 150$, $K_{i,FC} = 1000$. Since the tuning of the coefficient directly depends on measured error dynamics, each ANN provides the dynamics of tuning in a similar form. While in agent 1, the range of the coefficients is similar to each other, in agent 2, it is better to keep the coefficients different from each other.
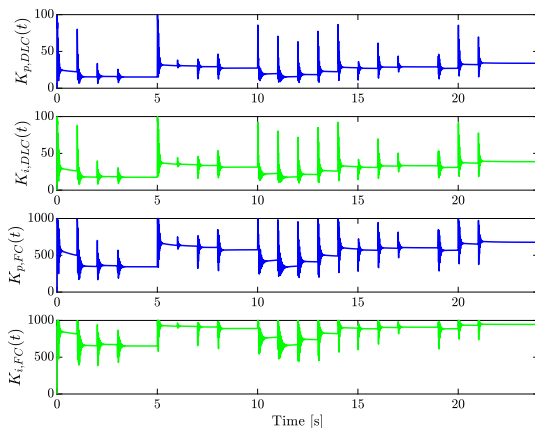

Fig. 6. Online tuning of two PI controllers DLC (top) and FC (bottom).

V. CONCLUSION

In this work, we proposed an online tuning method for the PI controller based on an ANN trained using the SRL algorithm to solve the FDC problem of the given isolated HMG. The results demonstrated that the proposed combination of PI controller and neural tuner and automatic tuning technique captures all the remarkable advantages of the controller, including robustness improvement and disturbance rejection, and gives rise to the generation of control laws that improve both transient and steady state cases.

REFERENCES

[1] M. J. Khan and M. T. Iqbal, "Dynamic modeling and simulation of a small wind-fuel cell hybrid energy system," Renew. Energ., vol. 30, pp. 421–439, 2005.

[2] O. C. Onar, M. Uzunoglu and M. S. Alam, "Dynamic modeling, design and simulation of a wind/fuel cell/ultra capacitor-based hybrid power generation system," J. Power Sources, vol. 161, pp. 707–722, 2006.

[3] M. Uzunoglu, O. C. Onar and M. S. Alam, "Modeling, control and simulation of a PV/FC/UC based hybrid power generation system for stand-alone applications," Renew. Energ., vol. 34, pp. 509–520, 2009.

[4] J. Ke, R. Xinbo, Y. Mengxiong and X. Min, "Power management for fuel-cell power system cold start," IEEE Trans. Power Electron., vol. 24, pp. 2391–2395, 2009.

[5] J. Ke, R. Xinbo, Y. Mengxiong and X. Min, "Comparative study of fuel-cell vehicle hybridization with battery or supercapacitor storage device," IEEE Trans. Veh. Technol., vol. 58, pp. 3892–3904, 2009.

[6] P. Thounthong, V. Chunkag, P. Sethakul, B. Davat and M. Hinaje, "Power management for fuel-cell power system cold start," IEEE Trans. Power Electron., vol. 24, pp. 2391–2395, 2009.

[7] M. Uzunoglu and M. S. Alam, "Modeling and analysis of an FC/UC hybrid vehicular power system using a novel wavelet based load sharing algorithm," IEEE Trans. Energy Convers., vol. 23, pp. 263–272, 2008.

[8] L. D. Jing and W. Li, "Small signal stability analysis of an autonomous hybrid renewable energy power generation/energy storage system part I: Time-domain simulations," IEEE Trans. Energy Convers., vol. 23, pp. 311–320, 2008.

[9] A. Doolla and T. S. Bhatti, "Load frequency control of an isolated small-hydro power plant with reduced dump," IEEE Trans. Energy Convers., vol. 21, pp. 1912–1919, 2006.

[10] M. Nayeripour, M. Hoseintabar and T. Niknam, "Frequency deviation control by coordination control of FC and double-layer capacitor in an autonomous hybrid renewable energy power generation system," Renew. Energ., vol. 36, pp. 1741–1746, 2011.

[11] C. D. Das, N. Sinha and A. K. Roy, "GA based frequency controller for solar thermal-diesel-wind hybrid energy generation/energy storage system," Int. J. Elec. Power, vol. 43, pp. 262–279, 2012.

[12] K. Nosrati, H. R. Mansouri and H. Saboori, "Fractional-order PID controller design of frequency deviation in a hybrid renewable energy generation and storage system," CIRED - Open Access Proc. J., vol. 1, pp. 1148–1152, 2017.

[13] A. Rahman, L. C. Saikia and N. Sinha, "Automatic generation control of an interconnected two-area hybrid thermal system considering dish-stirling solar thermal and wind turbine system," Renew. Energ., vol. 105, pp. 41–54, 2017.

[14] J. Mudi, C. K. Shiva, B. Vedik and V. Mukherjee, "Frequency stabilization of solar thermal-photovoltaic hybrid renewable power generation using energy storage devices," Iran. J. Sci. Technol. - Trans. Electr. Eng., vol. 45, pp. 597–617, 2021.

[15] B. Vedik, R. Kumar, R. Deshmukh, S. Verma and C. K. Shiva, "Renewable energy-based load frequency stabilization of interconnected power systems using quasi-oppositional dragonfly algorithm," J. Control Autom. Electr. Syst., vol. 32, pp. 227–243, 2021.

[16] V. Skiparev, J. Belikov and E. Petlenkov, "Reinforcement Learning based Approach for Virtual Inertia Control in Microgrids with Renewable Energy Sources," IEEE PES Innovative Smart Grid Technologies Conference Europe, pp. 1020–1024, 2020.

Appendix 6

IV

V. Skiparev, K. Nosrati, A. Tepljakov, E. Petlenkov, Y. Levron, J. Belikov, and J. M. Guerrero. Virtual inertia control of an isolated microgrid using NN-VFOPID controller: A self-tuning approach. *IEEE Transactions on Sustainable Energy*, 2023

Virtual Inertia Control of Isolated Microgrids Using an NN-Based VFOPID Controller

Vjatseslav Skiparev , *Student Member, IEEE*, Komeil Nosrati , *Member, IEEE*,
Aleksei Tepljakov , *Senior Member, IEEE*, Eduard Petlenkov, Yoash Levron , *Senior Member, IEEE*,
Juri Belikov , *Senior Member, IEEE*, and Josep M. Guerrero , *Fellow, IEEE*

Abstract—Reduction in system inertia and maintaining the frequency at the nominal value is a staple of today's and future power systems since their operation, stability, and resiliency are degraded by frequency oscillation and cascading failures. Consequently, designing a stable, scalable, and robust virtual inertia control system is highly relevant to skillfully diminishing the deviations during major contingencies. Therefore, considering the potential problems in predesigned nonflexible control systems with offline tuning techniques, we propose a variable fractional-order PID controller for virtual inertia control applications, which is tuned online using a modified neural network-based algorithm. The new proposed tuner algorithm is trained using a deep reinforcement learning strategy with a simplified deep deterministic policy gradient, which considers microgrid uncertainties. Compared with existing methods, all the tuning knobs of the discrete type and fully tunable variable FOPID controller (for both gain and order) can be captured based on the proposed hybrid algorithm, which inherits features from both classical and advanced techniques. To demonstrate the effectiveness of the training of the proposed controller, a comparative analysis with the standard FOPID and PID controllers is given under three different scenarios with a smooth (dis)connection of renewable energy sources and loads.

Index Terms—Microgrid, virtual inertia control, deep reinforcement learning, variable FOPID, neural networks, renewable energy.

NOMENCLATURE

ACE	Area control error.
CDM	Coefficient diagram method.
DG	Distribution grid.
DRL	Deep reinforcement learning.
ESS	Energy storage system.
FLC	Fuzzy logic control.
FOMCON	Fractional-order modeling and control.
FOPID	Fractional-order PID.
GRC	Generation rate constraint.
MG	Microgrid.
MPC	Model predictive control.
NN	Neural network.
PCL	Primary control loop.
PID	Proportional-integral-derivative.
PV	Photovoltaic.
RES	Renewable energy source.
SCL	Secondary control loop.
SDDPG	Simplified deep deterministic policy gradient.
TPP	Thermal power plant.
VFOPID	Variable FOPID.
VIC	Virtual inertia control.
VO-FD	Variable order fractional derivative.
VO-FI	Variable order fractional integral.
WTG	Wind turbine generator.

I. INTRODUCTION

AUTONOMOUS MGs are an increasing trend in power systems, where the high RES penetration is a significant challenge for the control of power supply stability, especially when the system is isolated from the grid [1], [2], [3]. Virtual inertia emulation is a solution that helps balance power systems with high-order dynamics [4]. This specific part of the virtual synchronous generator uses ESSs to compensate for the lack of rotational inertia and decrease the jeopardizing influence of RESs [5]. The parameters of these systems can be manipulated to enhance the dynamic system response. As a result, the concept of VIC with a suitable strategy is critical for MG operation [6].

To improve the stability of low-inertia MGs, different control structures, such as central and decentralized control, are used. While all control actions are made by a central unit or controllable distributed generations in the former structure, the latter offers local-based robust VIC techniques considering the increased uncertainties and nonlinearity contributed by the integration of renewable energy and distributed generation [7],

[8] to obtain more system efficiency and dynamic robustness [9]. This concept was first implemented in [10] using converters, and many other strategies have been developed in the literature since, with a special look at the control algorithm—as the core of the system—based on decentralized approaches. Several approaches have been proposed in this regard to solve the VIC problem with an adaptable virtual inertia constant, which is an important factor in emulating inertia power into the MGs, including optimal control [11], [12], H-infinity control [13], [14], CDM [15], integer and FOPID controllers [16], [17], [18], [19], [20], [21], [22], FLC [23], [24], DRL method [25], [26], and MPC [27], [28]. These approaches are frequently used as a reference for comparing the performance of VIC and can be divided into three main categories [2]:

- *Classical algorithms:* Most standard systems use classical algorithms such as H-infinity control [13], [14], CDM [15], and PID or FOPID-based controllers [16], [17], [18], [19], [20] for inertia emulation. Classical optimization, simplicity, and high robustness are three major benefits offered by these algorithms.
- *Advanced algorithms:* With FLC [23], [24], a self-adaptive system frequency stabilization, via manual design and without any on/offline optimization, was developed against the high amplitude fluctuations. For an optimal tradeoff between the transient frequency regulation and the respective control effort, the DRL method was applied to design an optimal controller [25], [26]. The major features of these algorithms are adaptation to uncertain conditions, online learning, and complexity, which is effective in multiloop control and adaptation to process dynamics but requires notable computational power.
- *Hybrid algorithms:* The algorithms in this category inherit features from the first two categories and can be used effectively to improve the system performance in future practical applications. The MPC [27], [28] and evolutionary-based optimization [19], [21], [22] of the classical controllers, associated with a data-driven approach, are two examples that cannot be classified into either of the abovementioned categories.

Recently, tuning-based FOPID and PID controllers were applied to the automatic load-frequency control of an islanded MG in a hybrid strategy [19], [21], [22]. Compared to other control techniques and according to the field researcher highlights, FOPID and PID controllers can properly generate the virtual inertia constant, which is the crucial term in emulating additional inertia power and was considered fixed before in control practice, which can imitate effective inertia power, resulting in improvement of the control system robustness against system perturbations and contribution to disturbance rejection control. By optimizing the controller parameters, it is possible to achieve not only a sufficient virtual inertia constant in relation to the additional power but also a stable grid operation. However, all previous studies attempted to achieve a higher degree of frequency control by adjusting parameters using traditional and intelligent rules in an offline approach. In these works, particularly for fractional cases, using the FOPID can bind a control engineer to manipulate all of the features of this controller,

such as the use of integration and derivative actions, toward a well-performing frequency control [29], [30], [31]. The online tuned PID controller from [32], applied to the superconducting transformer system, is based on the radial basis activation function and supervised learning optimization method, which can only find the local minimum.

To overcome the aforementioned control problems in MG related to the frequency deviation and the application of VIC, such as recovery time and instability, the selection of a proper control is very important. The online and automatic tuning-based FOPID controller, which captures all the remarkable advantages of the controller, including robustness improvement, disturbance rejection, and its contributions to time delay systems, is applied to generate the proper virtual inertia constant in a hybrid strategy. We aim to synthesize a new decentralized and flexible VFOPID to evaluate the virtual inertia power under different levels of renewable energy penetration and load disturbances. Therefore, we join the current trends in VIC modeling and propose using an additional NN-based VFOPID controller, in which the NN-based tuner is trained by the DRL algorithm. This hybrid strategy, including the FOPID controller, can improve the robustness of the system and, with its additional features and "tuning knobs," give rise to generating control laws that lead to superior frequency support loop performance. The application of the NN-based tuning method here is due to its learning ability and adaptability with parameter variation in the complex MG environment under intermittent dynamics. This technique was previously used to adjust the controller's parameters in an offline mode [33], [34] for frequency load control in hybrid MGs, resulting in nonflexible approaches that suffer from "dynamic adaptation".

The research motivation here addresses the frequency deviation control problem and the VIC application of islanded MGs based on an NN-based VFOPID controller associated with DRL and pursues the following objectives:

- Present an SDDPG with a semistochastic optimization algorithm that searches for optimal NN weights;
- Create an NN-based VFOPID controller trained using the proposed algorithm;
- Develop an online decentralized control strategy based on the designed automatically tuned VFOPID controller, including tuning variable order coefficients μ and λ;
- Develop an intelligent control policy that will fully utilize the potential of an ESS and provide improved frequency support; and
- Organize various smoothed connection-based scenarios involving RESs and loads.

It is worth mentioning that the proposed algorithm performs the additional virtual inertia emulation control using the scheme from [4], [35]. Here, and unlike the works [19], [21], [22], we introduce an NN as an online tuner for VFOPID controller coefficients. In contrast to [32], the DRL-based strategy together with the SDDPG policy technique is applied here to find the optimal weights of an NN-based VFOPID controller on a global scale. This policy with an extended reward system meets the requirements of the proposed hybrid frequency and ESS control. We create special rewarding rules to train the RL agent resulting
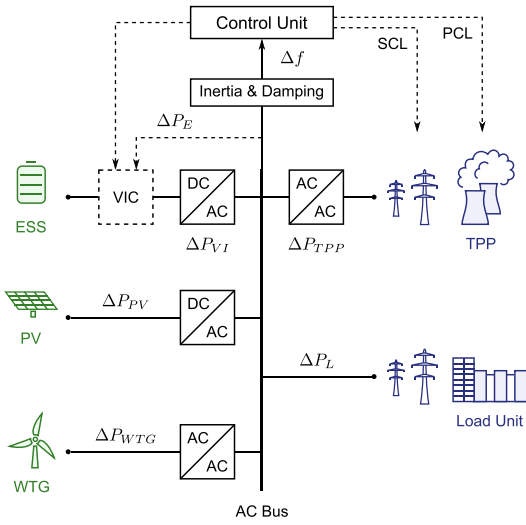
Fig. 1.　Schematic of the isolated MG with VIC.

Fig. 2.　Scheme of the VIC system, PCL, and SCL.

in a decrease in energy storage system power losses and a possible solution to balance energy cost and performance in future works. Furthermore, by providing a desirable performance related to frequency support problems, our method can provide many benefits to MG frequency support operations with high levels of renewable energy. We discover how all the tuning knobs of the FOPID controller can be captured by designing a self-tuning technique, resulting in robustness improvement, fast recovery time, and frequency stability enhancement.

II. PROBLEM STATEMENT

A. Structure of an Isolated MG

The studied MG presents a simplified version of a real AC-DC system without power convergence and is adopted from several recent publications [4], [13], [14], [15], [19], [25], depicted in Fig. 1. The addressed architecture includes 15 MW simplified domestic loads with a function for a smooth start, a 12 MW AC source (TPP), a 16 MW installed capacity of RESs (a WTG and a PV system), and an ESS with an extended VIC. The TPP is a major power producer composed of a governor with a GRC and a turbine with an applied frequency rate limiter, which restricts valve opening and closing actions (V_U, V_L) (see Fig. 2). Due to the restriction in the rate of change in the TPP unit, its physical dynamics are employed using GRC to regulate the turbine speed. Additionally, the delay in the speed governor is another significant constraint that necessitates the use of a dead band component for frequency regulation under various disturbances. Due to the small size of the turbine, the dead band saturation governor has a constant ± 0.12 p.u. MW/min for nonreheat generating units.

To control the frequency deviation Δf and preserve MG stability under disturbances, three main frequency control schemes
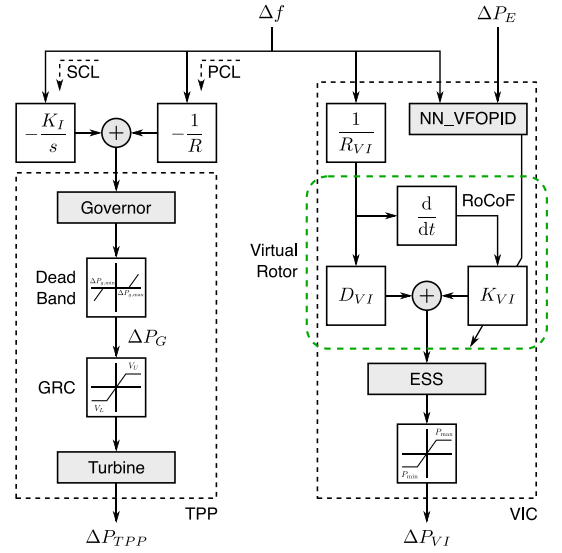
are employed: PCL, SCL, and VIC. In this hierarchical architecture, the PCL and SCL are responsible for balancing and restoring the system frequency, respectively. These two loops are applied to the TPP governor to generate power from the turbine system as

$$\Delta P_G = \frac{1}{1 + sT_G} \left(\Delta P_{SCL} - \Delta P_{PCL} \right), \qquad (1)$$

in which $\Delta P_{PCL} = R^{-1} \Delta f$ and $\Delta P_{SCL} = s^{-1} - K_I \Delta f$ are the control and ACE action changes from PCL and SCL, respectively, R is the droop constant, and K_I is the integral controller gain. By defining the load damping coefficient D and inertia constant H, the frequency deviation Δf can be represented as

$$\Delta f = \frac{1}{2Hs + D} \Delta P_E, \qquad (2)$$

where ΔP_E is the general power deviation resulting from all power sources and loads and can be calculated as

$$\Delta P_E = \Delta P_{TPP} + \Delta P_{WTG} + \Delta P_{PV} + \Delta P_{VI} - \Delta P_L, \quad (3)$$

with

$$\Delta P_{TPP} = \frac{1}{1 + sT_t} \Delta P_G, \qquad (4)$$

and $\Delta P_L = \Delta P_{RL} + \Delta P_{IL}$ as the load power changes, respectively. To make the system more feasible, the renewable energy and loads do not participate in frequency management and are considered system uncertainties. Hence, the following simplified models of the renewable energy power changes are sufficiently accurate for our analysis:

$$\Delta P_{RES} = \Delta P_{WTG} + \Delta P_{PV}$$

$$= \frac{1}{1 + sT_{WTG}} \Delta P_{wind} + \frac{1}{1 + sT_{PV}} \Delta P_{solar}. \quad (5)$$

TABLE I
PARAMETERS OF THE MG

Parameter	Physical meaning	Nominal value	Unit
ΔP_{TPP}	TPP power changes	–	p.u.
ΔP_E	General power deviations	–	p.u.
H	System inertia	0.083	p.u. s
D	Damping coefficient	0.015	p.u./Hz
ΔP_{VI}	Virtual inertia power changes	–	p.u.
K_{VI}	Virtual inertia constant	–	p.u. s
D_{VI}	Virtual damping coefficient	0.3	p.u./Hz
R_{VI}	Virtual inertia droop coefficient	2.7	Hz/p.u.
R	Droop coefficient	2.4	Hz/p.u.
T_{VI}	Virtual inertia time constant	10	s
T_t	Turbine time constant	0.4	s
ΔP_G	Governor power changes	–	p.u.
T_G	Governor time constant	0.1	s
ΔP_{SCL}	ACE action changes of SCL	–	p.u.
ΔP_{PCL}	Control changes of PCL	–	p.u.
K_I	Integral controller gain	0.075	s
ΔP_{RES}	RES power changes	–	p.u.
ΔP_{WTG}	WTG power changes	–	p.u.
ΔP_{wind}	Initial wind power variation	–	p.u.
T_{WTG}	WTG time constant	1.85	s
ΔP_{PV}	PV power changes	–	p.u.
ΔP_{solar}	Initial solar power variation	–	p.u.
T_{PV}	PV time constant	1.5	s
ΔP_L	Load power changes	–	p.u.
ΔP_{RL}	Residential loads variations	–	p.u.
ΔP_{IL}	Industrial loads variations in	–	p.u.
V_U, V_L	Governor valve limiter	±0.5	p.u.
$\Delta P_{g,\min}^{\max}$	Governor dead band limits	±0.12	p.u. MW/min
$\Delta P_{VI,\min}^{\max}$	Virtual inertia valve limiter	±0.1	p.u.

The given parameters in the equations are summarized in Table I.

B. VIC Modeling

The applied decentralized controller scheme is adapted from [4] and is illustrated in Fig. 2. The scheme describes the behavior of an electronics-based power inverter applied to the inertia emulation. The virtual inertia constant K_{VI} is usually defined as

$$K_{VI} = \frac{2HP_{Inv}}{f_0}, \qquad (6)$$

where f_0 is the nominal frequency, P_{Inv} is the power of the inverter, and H is the calculated inertia [15], [19], [35]. To mimic the necessary inertia power, K_{VI} must be able to adapt to a variety of disturbances without jeopardizing the system stability. The original scheme with a directly adjustable junction K_{VI} by an additional controller was addressed in several recent publications [14], [15], [28]. The extended version includes the damping constant D_{VI} and junction variable K_{VI} adjusted by an additional controller. In this work, the proposed additional controller is affected by input signals Δf and ΔP_E, in which

the latter is an estimated power system measurement value due to its complexity [36]. Furthermore, as shown in Fig. 2, the scheme employs PCL and SCL with Δf as an input signal, which reduces the influence of noise and controls the active power in relation to frequency deviation. The dynamic equation of the power changes ΔP_{VI} can be given by

$$\Delta P_{VI} = \frac{sK_{VI} + D_{VI}}{1 + sT_{VI}}\left(\frac{\Delta f(s)}{R_{VI}}\right). \qquad (7)$$

The rate of change in frequency (RoCoF) can be expressed as

$$\text{RoCoF} = \frac{\mathrm{d}\Delta f}{\mathrm{d}t}. \qquad (8)$$

In the proposed dynamic structure, the derivative control approach is the fundamental behavior behind the VIC, where it can compute the RoCoF to adjust an extra power to a given point during RES penetration and contingency [37] and gives the optimum frequency response due to its impact on the system inertia and damping [38], [39]. In this regard and to reduce power oscillations associated with better inertia support, the proposed VFOPID controller is considered to manipulate the virtual inertia unit in choosing the proper $K_{VI}(t)$ in each time step t for emulating sufficient inertia power against various changes. To achieve an optimal K_{VI}, we propose a methodology and design an algorithm for the VIC-based VFOPID feedback control technique to calculate a reasonable approximation in estimating a suitable value for K_{VI} and provide the additional power transfer ΔP_{VI} when Δf and RoCoF exceed defined limits.

III. PROPOSED SOLUTION

The proposed methodology is based on the reinforcement learning architecture neural actor-critic [40], [41], which is used to train an NN-based online tuner of the VFOPID controller coefficients. The tuner is implemented as an actor whose interactions are observed by the critic according to the SDDPG policy estimation of the MG state.

A. VFOPID Controller

The original FOPID controller was first introduced in its parallel form described as [42]:

$$u(t) = K_p e(t) + K_i D^{-\lambda} e(t) + K_d D^{\mu} e(t), \qquad (9)$$

where D^{α} is the fractional differential ($\alpha > 0$) (or integral ($\alpha < 0$)) operator, μ and λ are the positive integration and differentiator orders, respectively, $u(t)$ is the control signal, $e(t)$ is the error signal, which is equal to $\Delta f(t)$, and K_p, K_i, and K_d are the proportional, integration, and derivative constant gains, respectively. The controller (9) has more tuning freedom that leads to making the plant stable under control and fulfilling intricate control performance requirements that are not in the scope of classical controllers. However, because this controller binds a control engineer to manipulate all of the features, the VFOPID controller with five variable parameters can significantly improve the controlled system performance due to its greater flexibility. Referring to Fig. 3, this means that it is
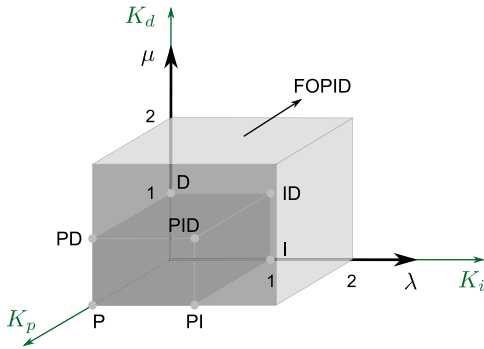
Fig. 3. Relationships between the VFOPID controller and its classical variations.

Fig. 4. Scheme of the VFOPID NN-based tuner (actor NN).

possible to move continuously not only in the PID plane instead of jumping between the fixed points but we can also explore for the desired controller values in a space inside the cube and between the eight vertices.

Both integration and differential functions affect the steady-state process and the dynamic characteristics when both the associated gains and orders of each function vary:

- *Integral action:* This action is dependent on the gain K_i and the order λ. While the oversize of K_i will make a system more unstable, its small size will force the system to diverge from its ideal dynamic performance. Additionally, for a small λ, the frequency band is wide, and the system is steady with rapid response and static error. In contrast, an oversized λ will degrade the system stability with the increase in the overshoot, rise, and settling times.
- *Derivative action:* The gain K_p has no effect on the steady-state error but can improve the dynamic characteristics. For a small K_p, the overshoot and settling time will increase, while for a large gain, the system noise may increase, and the system performance will degrade. Finally, the order μ can improve the response accuracy and steady-state error when it is relatively small. However, by increasing μ, the overshoot and settling time decrease, and the closed-loop system stability will degrade.

It is necessary to keep both the gains and orders in a suitable range at any moment to maintain a satisfactory control performance for a target system. To take advantage of this controller flexibility and promote controlled system performance simultaneously, the VFOPID needs to be considered as a target controller in the regulation process. The concept of VFOPID can be given as [43]:

$$u(t) = K_p(t)e(t) + K_i(t)D^{-\lambda(t)}e(t) + K_d(t)D_t^{\mu(t)}e(t), \tag{10}$$

where $D^{-\lambda(t)}$ and $D_t^{\mu(t)}$ are the variable order fractional integral and derivative, respectively.

The implementation of (10) with a floating point requires a powerful computational device [44]. Therefore, we propose using a parallel connection of FOMCON library blocks from [29] and provide a real-time switch between fractional-order
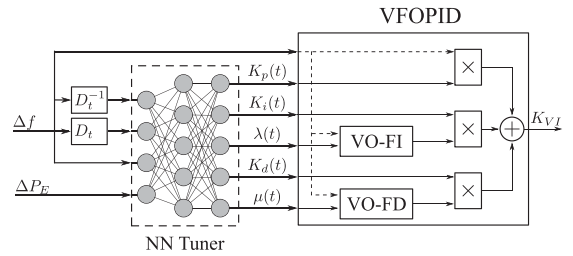
constants with a resolution of 0.1. In conjunction with switching objective functions and arising from an input signal derived from the NN tuner system, these functions attempt to switch between different fractional integration and differentiator operators and implement a VFOPID controller according to the tuned parameters. By doing so, the fractional-order tuning and controller gains can be performed online.

B. NN-Based Tuning of the VFOPID Controller

We apply the multiple input multiple output type of NN as the tuner of the VFOPID controller. The major task here for actor NN (i.e., tuner) is to search for an optimal combination of these coefficients considering MG disturbances. Fig. 4 illustrates the proposed architecture of the NN-based tuner, where output neurons are five controller parameters: K_p, K_i, K_d, λ, and μ. Since the values of each output neuron should be nonnegative, we designed the network to keep outputs for the fractional orders and gains of the controller in the range [0,2] defined by the equation

$$y(t) = \frac{2}{1 + e^{-2x(t)}}. \tag{11}$$

In fact, the orders λ and μ are very sensitive to any changes and are computationally expensive. Therefore, we organized the change in the integral and derivative parameters with the discrete step of 0.1 by using an automatic switch between predefined series of FOMCON blocks implemented in the MATLAB/Simulink environment, where each block has a frequency range $[0.001, 1000]$ and approximation order 3. In the proposed method, tuning the VFOPID parameters depends on the magnitude of Δf being proportional to the power variation ΔP_E, which is driven to the NN tuner (see Fig. 4), and is then transformed by the $tansig$ activation function illustrated in (11) to give the tuned parameters as the given trajectories depicted in Fig. 9. The output value of every coefficient depends on the gain provided by every weight considered in the NN tuner.

C. DRL-Based VFOPID Controller

The key advantage of reinforcement learning is the direct interaction of NN with a controllable environment and the application of nature-inspired reward-punishment-based learning, by which it can theoretically find a global minimum after training procedure [41], [45]. Usually, reinforcement learning is applied in applications where it is necessary to train NN for tasks in
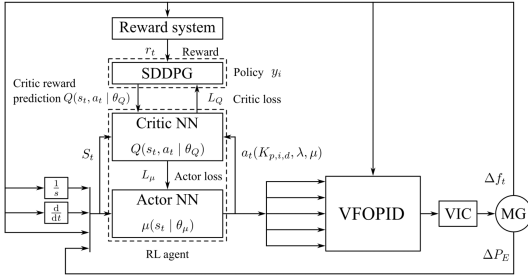
Fig. 5. Scheme of the proposed DRL-based VFOPID for VIC in an MG.

Algorithm 1: SDDPG algorithm with semistochastic method.

1:     Initialize critic $Q(s, a \mid \theta_Q)$ actor $\mu(s \mid \theta_\mu)$ networks
2:     **for** $t = 1$ to N **do**
3:        Initialize random weights of actor θ_μ.
4:        Execute action a_t, observe reward r_t at each step
5:        **if** $r_{avg} > r_{avg\_max} = \max(\frac{1}{N} \sum_i r_i(s_i, a_i))$ **then**
6:           Save weights θ_Q and θ_μ with best r_{avg}.
7:        **end if**
8:     **end for**
9:     Initialize random weights of critic θ_Q.
10:    Define the learning rate of critic α_Q and actor α_μ networks.
11:    Define the smooth factor constant $\tau \in [0, 1]$.
12:    **for** $episode = 1$ to M **do**
13:       **for** $t = 1$ to T **do**
14:          Receive initial process observation as state s_t.
15:          Select action $a_t = \mu(s_t \mid \theta_\mu)$ according to current reward prediction $Q(s_i, a_i \mid \theta_Q)$.
16:          Execute action a_t, observe reward r_t and future state s_{t+1}.
17:          Select the best critic reward prediction weights $Q'(s_{i+1}, \mu'(s_{i+1}|\theta_{\mu'})$ according to the explored r_{avg_max}.
18:          Set the state-action-value function of DDPG policy $y_i = r_i + \gamma Q'(s_{i+1}, \mu'(s_{i+1}|\theta_{\mu'}) \mid \theta_{Q'})$.
19:          Update the critic by minimizing the loss: $L_Q = \frac{1}{N} \sum_i (y_i - Q(s_i, a_i \mid \theta_Q))^2$.
20:          Update actor policy using the sampled gradient of critic: $L_\mu = \frac{1}{N} \sum_i (Q(s, a \mid \theta_Q)|_{s=s_i, \alpha=\mu(s_i)})^2$
21:          Update target networks:
22:     $\theta^{Q'} \leftarrow \tau \theta^Q + (1-\tau) \theta^{Q'} \alpha_Q L_Q$,
23:     $\theta^{\mu'} \leftarrow \tau \theta^\mu + (1-\tau) \theta^{\mu'} \alpha_\mu L_\mu$.
24:       **end for**
25:    **end for**

which the correct interaction of an autonomous system with the environment is critically important [40], [41]. The general structure of the proposed controller is presented in Fig. 5. How to properly interact with VFOPID and MG by "double feedback," i.e., from the reward system and from the process directly, is seen in the presented architecture of neural actor-critic studies. The first operates in offline mode and the second in online mode, which can be summarized as:

- Reward or punish the agent by directing the measured error Δf into the block "reward system".
- Formulate the observation for the agent using the measured variables Δf and ΔP_E, which can be considered as the control error and preerror, respectively.

1) Neural Actor-Critic: The neural actor-critic architecture requires synchronizing two NNs: the actor network $\mu(s \mid \theta_\mu)$ and the critic network $Q(s, a \mid \theta_Q)$. The actor network in this architecture is an agent that interacts directly with the environment, while the critic network $Q(s, a \mid \theta_Q)$ observes any changes in the environment state s_t, and the reaction of the actor a_t corrects its interaction according to the defined policy (e.g., dynamic programming, Monte-Carlo policy gradient, stochastic policy gradient, Q-learning, DDPG). This network uses the loss function L_μ to correct actor weights to find a potential right estimation of the state-action-value function $Q(s, a)$ with a prediction of possible average reward r_{avg} per epoch. To increase the quality of the prediction, we store the best prediction as $Q'(s_{i+1}, \mu'(s_{i+1}|\theta_{\mu'}) \mid \theta_{Q'})$ applied for the state-action-value function.

2) SDDPG with Semistochastic Approach: Original DDPG is a model-free algorithm and is designed for low-dimensional continuous action space tasks [40]. Deep Q-learning [41] and deterministic policy gradient [40], [46] methods are combined in the DDPG optimization technique, which inherits the neural actor-critic architecture. Unlike the original DDPG, the modified version has a simplified loss function for the actor L_μ, and learning is performed without the replay buffer and noise exploration. The optimization principle is based on minimizing the difference between the target action-value function y_i and the critic network reward prediction $Q(s_i, a_i \mid \theta_Q)$. For each actor NN $\mu(s \mid \theta_\mu)$, the decision is defined as a_t (i.e., $K_p(t)$, $K_i(t)$, $K_d(t)$, $\lambda(t)$, $\mu(t)$) per step i in MG state s_t (i.e., $\Delta f_t, \Delta f_t/s, \mathrm{d}\Delta f/\mathrm{d}t, \Delta P_E$). The SDDPG algorithm seeks to maximize the average reward r_t per training episode M by minimizing the critic L_Q and actor L_μ

loss functions [40], [47]. Unlike [47], we extend the algorithm by combining it with the stochastic approach, in which we initialize random weights N times and choose parameters with the highest average reward r_{avg_max} after beginning training with SDDPG. This approach can help increase the speed of the optimization process. The pseudocode of the proposed method is shown in Algorithm 1, which summarizes the SDDPG algorithm associated with the semistochastic approach.

3) Agent Reward System: To provide agent learning, we use the reward/punishment system proposed in [25], [26]. The reward system is organized in such a way that the measured frequency deviation is converted to reward/punishment $r_t\pm$. To provide the instructions according to the stability criteria of error signal Δf in the ± 0.1 Hz band [2], [4] and inertia injection ΔP_{VI} in range ± 0.1 p.u., the regulation system is organized as follows: if $|\Delta f| < 0.05$ and if $|\Delta P_{VI}| < 0.075$, the agent receives a reward $r+$; otherwise, the system punishes $r-$ after each performed action. To provide a reasonable reward for the agent we separated the rewarding approach for
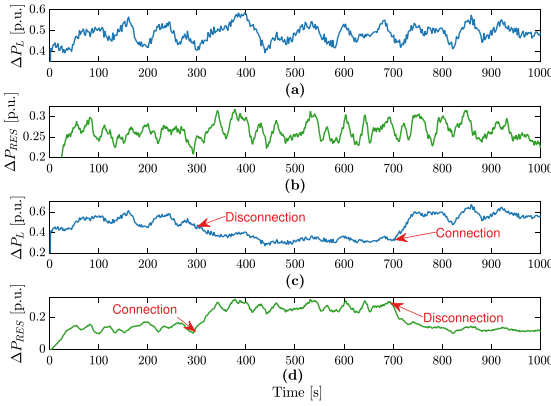
Fig. 6. Data used in the considered scenarios. Cases (a, b) apply to Scenario 1 (Nominal), (a, d) to Scenario 2 (RESs), and (b, c) to Scenario 3 (Loads) with arrows indicating (dis)connection moments.

each task, the frequency support reward task is limited to the range $u \in \{0.05, \ldots, 2\}$, and the punishment is unlimited and multiplied by 2. In the inertia injection task, control the reward multiplied by 0.1 and punishment by 0.5. The following equation presents the mathematical expression for the designed rules:

$$r_t = \begin{cases} \frac{1}{0.5 + |\Delta f|}, & \text{if } |\Delta f| < 0.05, \\ -2|\Delta f|, & \text{if } |\Delta f| > 0.05. \\ 0.1|\Delta P_{VI}|, & \text{if } |\Delta P_{VI}| < 0.075 \\ -0.5|\Delta P_{VI}|, & \text{if } |\Delta P_{VI}| > 0.075 \end{cases} \quad (12)$$

IV. NUMERICAL RESULTS

In this section, we analyze the performance of the proposed algorithm (NN-based VFOPID controller) and compare it with FOPID and PID. The controllers were carefully tuned, and the obtained parameters were $K_p = 50$, $K_i = 1.75$, $K_d = 2$, $\lambda = 1.25$, and $\mu = 1.75$ for the FOPID controller and $K_p = 50.5$, $K_i = 5.85$, and $K_d = 5.5$ for the PID controller. These parameters were obtained by heuristic search, where we attempted to find the optimal parameters for the studied system and provide a fair comparison with the proposed method. All simulations were performed in the MATLAB/Simulink environment, and the different parameters were previously described. In these experiments, we perform simulations for three different scenarios: the nominal case (Scenario 1), the case with (dis)connection of RESs (Scenario 2), and the case with smooth (dis)connection of loads (Scenario 3). Each scenario is simulated for nominal (100%) and decreased inertia (40%). Fig. 6 shows the dynamics of renewable energy and loads. Subplots (a) and (b) illustrate the data used in the nominal scenario, while subplots (c) and (d) show scenarios with loads and renewable energy (dis)connections. To replicate more natural power disturbances, smooth (dis)connections are provided.
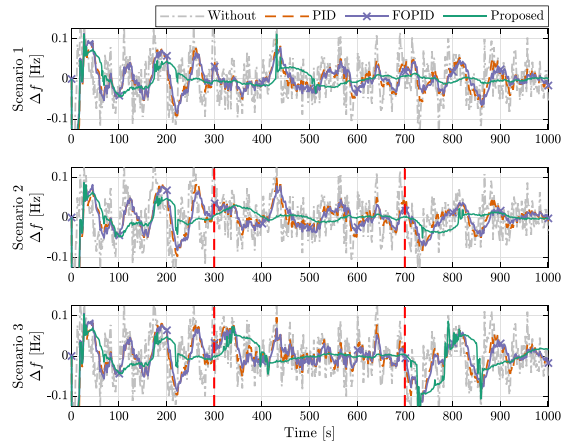
Fig. 7. Frequency deviation. Performance results of the proposed (green), FOPID (purple with crosses), PID (dashed orange) and without controller (dashed gray) for Scenarios 1–3 with 100% inertia. Red dashed lines indicate (dis)connection moments.
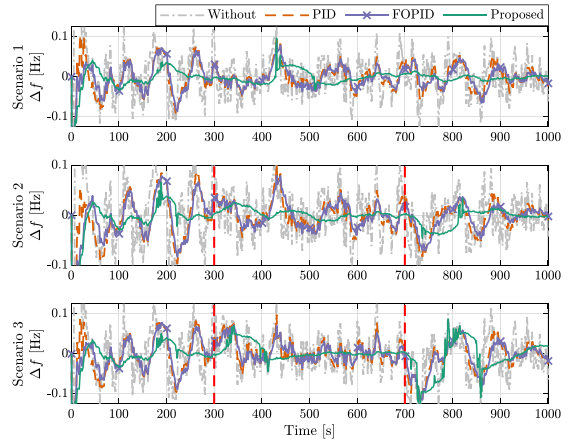
Fig. 8. Frequency deviation. Performance results of the proposed (green), FOPID (purple with crosses) and PID (dashed orange) and without controller (dashed gray) for Scenarios 1–3 with 40% inertia. Red dashed lines indicate (dis)connection moments.

A. Comparative Analysis

Here, we compare the mentioned algorithms in three different scenarios. We remove the first 100 samples to eliminate the effect of high oscillations appearing at the beginning of the simulations due to random initial conditions.

Scenario 1 (Nominal): Recall that in this case, we do not provide any (dis)connections. The top plots in Figs. 7 and 8 show the performance results of the three considered algorithms for the cases with 100% and 40% inertia levels, respectively. It can be observed that the proposed algorithm provides smooth frequency support with a deviation of less than ± 0.075 Hz. The other two algorithms, FOPID and PID, provide reasonable performance with some notable oscillatory behavior.

TABLE II
COMPARISON OF DIFFERENT CONTROLLERS USING THE IAE METRIC

Scenario	Inertia	Proposed	FOPID	PID	Without*
Scenario 1	100%	**10.28**	19.99	22.31	34.20
	40%	**9.05**	20	22.33	33.48
Scenario 2	100%	**11.10**	21.23	22.81	34.31
	40%	**9.18**	21.26	22.91	33.77
Scenario 3	100%	**19.50**	23.65	26.44	37.18
	40%	**18.14**	23.72	26.51	36.19

*Hereinafter, 'Without' means that no additional VIC is applied.

TABLE III
COMPARISON OF DIFFERENT CONTROLLERS USING THE RMSE METRIC

Scenario	Inertia	Proposed	FOPID	PID	Without
Scenario 1	100%	**0.0157**	0.0281	0.0311	0.0476
	40%	**0.0136**	0.0282	0.0311	0.0463
Scenario 2	100%	**0.0175**	0.0300	0.0328	0.0482
	40%	**0.0147**	0.0301	0.0330	0.0471
Scenario 3	100%	**0.0324**	0.0351	0.0379	0.0523
	40%	**0.0312**	0.0352	0.0380	0.0511

Scenario 2 (RESs (dis)connection): To compare the algorithms, an additional (dis)connection-based scenario is carried out, and the performance is evaluated. We organize a smooth connection of solar panels and the disconnection of wind turbines in time steps 300 and 700 seconds, respectively (see Fig. 6(d)). Figs. 7 and 8 depict a significant influence of the wind turbine connection on the performance of all control algorithms. However, the proposed strategy shows the best performance and again results in smooth transient behavior.

Scenario 3 (Loads (dis)connection): Here, we provide the test with the smooth connection and disconnection of residential loads. This scenario appears to be the most challenging for all considered algorithms, as depicted in Figs. 7 and 8. The proposed method shows the best performance, while the FOPID and PID controllers provide reasonable performance.

To quantify the effectiveness of all control algorithms, we use the integral of the absolute value of error (IAE) metric (i.e., total frequency deviation): $\text{IAE} = \int_0^T |\Delta f| \mathrm{d}t$, where T is the simulation time period. Then, the performance is evaluated using the absolute error $\text{MAE} = \sum_{i=0}^N |\Delta f_i|$, where N is the number of samples. In addition, we evaluate the performance using the root mean square error (RMSE) metric given by:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=0}^N (\Delta f_i)^2}. \tag{13}$$

Performance indices are calculated in the steady-state conditions (i.e., from 100 to 1,000 seconds) to eliminate the effect of high initial oscillations. All calculations are performed for three scenarios with both nominal (100%) and decreased inertia (40%). The results are summarized in Tables II–IV. It can be seen that the proposed NN-based VFOPID controller has better results in all cases.

TABLE IV
COMPARISON OF DIFFERENT CONTROLLERS USING THE MAE METRIC

Scenario	Inertia	Proposed	FOPID	PID	Without
Scenario 1	100%	**0.0114**	0.0222	0.0248	0.038
	40%	**0.0100**	0.0223	0.0248	0.0372
Scenario 2	100%	**0.0123**	0.0236	0.0253	0.0381
	40%	**0.0102**	0.0235	0.0254	0.0375
Scenario 3	100%	**0.0180**	0.0262	0.0293	0.0413
	40%	**0.0168**	0.0263	0.0294	0.0402

Fig. 9. Scenario 1 (Nominal). Evolution of gains and orders of the proposed controller for both nominal (100%, blue) and decreased inertia (40%, red dashed).

B. Online Tuning of the VFOPID

Figs. 9– 11 present the online optimization (tuning) results of all five VFOPID controller parameters. This is done for three different scenarios and two inertia cases, as detailed in the previous section.

Scenario 1 (Nominal): Fig. 9 depicts that the VFOPID coefficients have rather stable behavior, all parameters vary 500 seconds and later stabilize, and the influence of decreased inertia is relatively small.

Scenario 2 (RESs (dis)connection): Unlike the above case, Fig. 10 shows the output of the neural tuner, where the influence of renewable sources and decreased inertia on the VFOPID tuner is significant. The integral knob K_i has a notable increase after connecting the wind turbine from 300 to 400 seconds and decreases after disconnecting the solar panels in 700 seconds. A similar change has K_p, K_d, μ, and λ.

Scenario 3 (Loads (dis)connection): Similar to the previous cases, Fig. 11 shows a notable influence of load disconnection (at $t = 300$s) on parameters K_p, K_d, μ, and λ. The integral gain K_i has a notable peak from 320 to 400 seconds. Additionally, after reconnecting loads at the time step 700 s, the gains and orders change significantly and attempt to return to their nominal range variations. However, the influence of decreased inertia is still minimal.
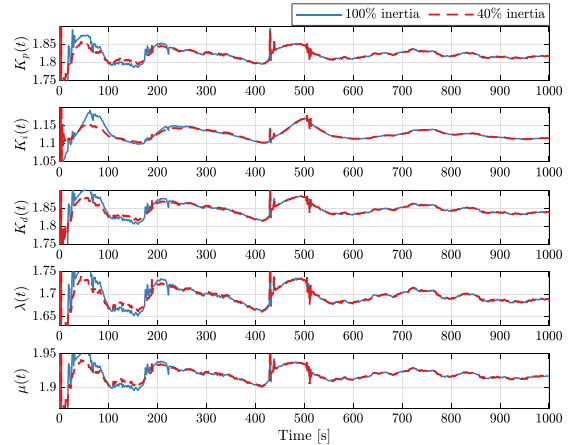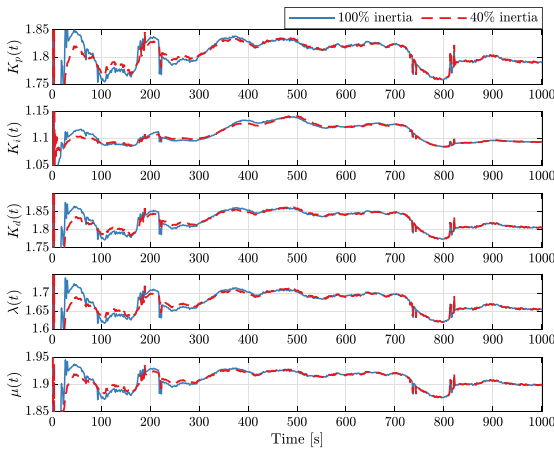
Fig. 10.   Scenario 2 (RESs (dis)connection). Evolution of gains and orders of the proposed controller for both nominal (100%, blue) and decreased inertia (40%, red dashed).
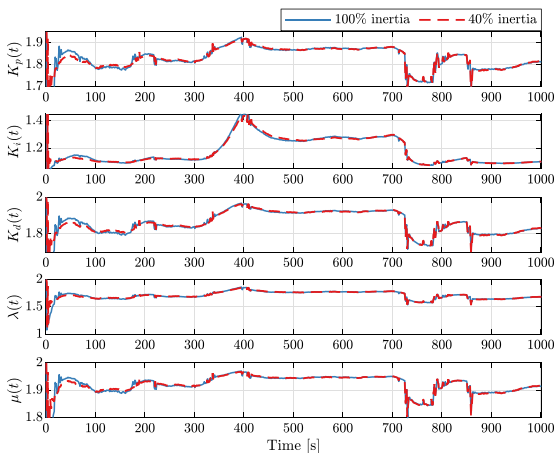

Fig. 11.   Scenario 3 (Loads (dis)connection). Evolution of gains and orders of the proposed controller for both nominal (100%, blue) and decreased inertia (40%, red dashed).

TABLE V
POWER BALANCE FOR DIFFERENT CONTROLLERS

	Inertia	Proposed	FOPID	PID	Without
Scenario 1	100%	0.1384	−1.0669	−1.9882	−0.0178
	40%	0.2098	−0.9745	−1.8653	−0.0182

C. Energy Storage Dynamics

Table V provides a criterion, calculated as $y = \int_0^N \Delta P_{VI} dt$, which summarizes the final balance between injected (positive) and absorbed (negative) power by the storage system in the nominal scenario. It can be seen that the proposed controller uses energy storage more effectively, as the final balance tends to be minimal in all cases. Fig. 12 illustrates that the proposed
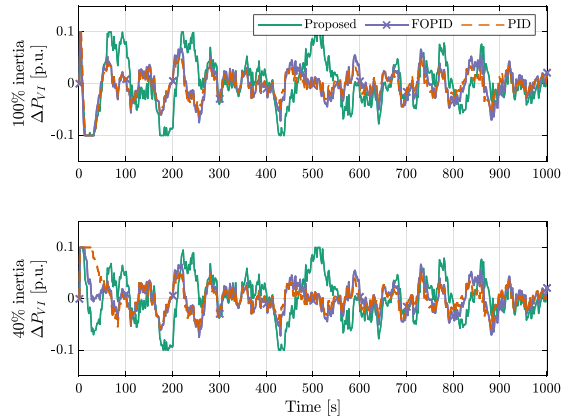

Fig. 12.   Power flow in ESS. Comparison of algorithms: the proposed (green), FOPID (purple with crosses), and PID (dashed orange) in the nominal scenario with 100% (top plot) and 40% inertia (bottom plot).

algorithm is forced to inject and store slightly more energy than other controllers; however, this happens only at certain time instances (approximately 180 and 420 s). This is done to guarantee smooth frequency support, which is encoded in the designed algorithm.

V. DISCUSSION

Integration of renewable energy into microgrid systems reduces the system inertia, which is the striking capability to sustain the frequency at its nominal value, resulting in a stable and resilient MG as well as the main grid. Renewable energy may increase the level of uncertainty during abnormal operations, introduce some technical implications to the VIC concept, and raise some questions about the sufficiency of traditional control approaches. The main question is what happens in a MG when there are natural changes and uncertainties, where traditional controllers cannot provide robust performance.

In response to this question, the present paper can be considered an effort to design a stable, scalable, and robust control system, which is highly relevant to skillfully diminishing the deviations during major contingencies. Although some recent studies have attempted to achieve a higher degree of frequency support by adjusting the parameters of well-known FOPID controllers using classical rules, offline tuning approaches can limit a control engineer's ability to manipulate all the features therein. Without intelligent control methods, it will be difficult to control the virtual inertia in modern MGs, which are continuously changing. For this purpose in real practice, the VIC system should be flexible and robust to enable a matching mechanism over a wide range of operations.

In this regard, we proposed an NN-based tuner for the VFOPID controller and both coefficients and fractional orders using the reinforcement learning strategy together with the SD-DPG optimization technique. Our method can offer many benefits for MG frequency support operations with renewable energy by providing desirable performance related to load-frequency

support problems. Additionally, we discovered how all the tuning knobs of such a controller can be captured by designing a self-tuning technique, resulting in robustness improvement, fast recovery time, and frequency stability enhancement.

VI. Conclusion

In this paper, we presented the novel combination of NN with a VFOPID controller, applied to robust virtual inertia emulation in isolated microgrids with high renewable energy penetration, where all tuned parameters (both gains and orders) of this controller were optimized by the semi-stochastic DRL algorithm with SDDPG policy. We proved that VFOPID controller fusion with the NN-based tuner is an effective strategy for virtual inertia emulation tasks. In contrast to other proposed algorithms, the NN-based VFOPID controller has active support from a machine learning algorithm that makes it self-adaptive to MG disturbances, resulting in smooth deviation but a more significant influence from decreased inertia. Compared to recent efforts, the major benefits of the proposed solution can be summarized as follows:

- An effective NN-tuning-based system for VFOPID controller coefficients uses the SRL strategy together with the SDDPG technique to find optimal weights for the controller.
- Fast training of the NN-based tuner by the proposed effective semi-stochastic RL.
- Capturing all features of the FOPID controller using a self-tuning technique, resulting in robustness improvement, fast recovery time, and frequency stability enhancement.
- Capability of the designed controller over a wide range of operating conditions is due to its flexibility in the use of integration and derivative actions toward a good performance in frequency support operations with high levels of renewable energy.

Further work is needed to address some technical issues for the VIC system in terms of algorithm performance evolution, including computational complexity reduction, accuracy enhancement, and robustness improvement. In this regard, implementing a real and high-resolution VFOPID controller associated with a constrained parameter search space based on a MG stability region derived by using the stability boundary locus method can be a future effective strategy for more complex systems such as multi-area interconnected MGs and power electronics. In addition, in an actual power system with high-order of complexity, the intermittent dynamics can affect the algorithm performance and routine process of the system. Therefore, developing a robust RL strategy, by upgrading the reward rules and integrating a digital protection system during the wrong training process, is also another actual control problem to be considered in the future.

References

[1] A. Ulbig, T. S. Borsche, and G. Andersson, "Impact of low rotational inertia on power system stability and operation," *IFAC Proc. Volumes*, vol. 47, no. 3, pp. 7290–7297, 2014.

[2] V. Skiparev, R. Machlev, N. R. Chowdhury, Y. Levron, E. Petlenkov, and J. Belikov, "Virtual inertia control methods in islanded microgrids," *Energies*, vol. 14, no. 6, 2021, Art. no. 1562.

[3] F. Dörfler, J. W. Simpson-Porco, and F. Bullo, "Breaking the hierarchy: Distributed control and economic optimality in microgrids," *IEEE Trans. Control Netw. Syst.*, vol. 3, no. 3, pp. 241–253, Sep. 2016.

[4] T. Kerdphol, F. S. Rahman, M. Watanabe, and Y. Mitani, *Virtual Inertia Synthesis and Control*. Berlin, Germany: Springer, 2021.

[5] B. Kroposki et al., "Achieving a 100% renewable grid: Operating electric power systems with extremely high levels of variable renewable energy," *IEEE Power Energy Mag.*, vol. 15, no. 2, pp. 61–73, Mar./Apr. 2017.

[6] F. Milano, F. Dörfler, G. Hug, D. J. Hill, and G. Verbic, "Foundations and challenges of low-inertia system (invited paper)," in *Proc. IEEE Power Syst. Comput. Conf.*, 2018, pp. 1–25.

[7] A. J. Veldhuis, M. Leach, and A. Yang, "The impact of increased decentralised generation on the reliability of an existing electricity network," *Appl. Energy*, vol. 215, pp. 479–502, 2018.

[8] N. Bazmohammadi et al., "Microgrid digital twins: Concepts, applications, and future trends," *IEEE Access*, vol. 10, pp. 2284–2302, 2022.

[9] H. Bevrani, F. Habibi, P. Babahajyani, M. Watanabe, and Y. Mitani, "Intelligent frequency control in an AC microgrid: Online PSO-based fuzzy tuning approach," *IEEE Trans. Smart Grid*, vol. 3, no. 4, pp. 1935–1944, Dec. 2012.

[10] H.-P. Beck and R. Hesse, "Virtual synchronous machine," in *Proc. IEEE 9th Int. Conf. Elect. Power Qual. Utilisation*, 2007, pp. 1–6.

[11] M. Datta, H. Ishikawa, H. Naitoh, and T. Senjyu, "Frequency control improvement in a PV-diesel hybrid power system with a virtual inertia controller," in *Proc. IEEE 7th Conf. Ind. Electron. Appl.*, 2012, pp. 1167–1172.

[12] M. Torres and L. A. Lopes, "An optimal virtual inertia controller to support frequency regulation in autonomous diesel power systems with high penetration of renewables," in *Proc. Int. Conf. Renewable Energies Power Qual.*, 2011, pp. 13–15.

[13] T. Kerdphol, F. S. Rahman, Y. Mitani, M. Watanabe, and S. Kufeoglu, "Robust virtual inertia control of an islanded microgrid considering high penetration of renewable energy," *IEEE Access*, vol. 6, pp. 625–636, 2018.

[14] T. Kerdphol, F. S. Rahman, M. Watanabe, and Y. Mitani, "Robust virtual inertia control of a low inertia microgrid considering frequency measurement effects," *IEEE Access*, vol. 7, pp. 57550–57560, 2019.

[15] H. Ali et al., "A new frequency control strategy in an islanded microgrid using virtual inertia control-based coefficient diagram method," *IEEE Access*, vol. 7, pp. 16979–16990, 2019.

[16] P. K. Mohanty, B. K. Sahu, T. K. Pati, S. Panda, and S. K. Kar, "Design and analysis of fuzzy PID controller with derivative filter for AGC in multi-area interconnected power system," *IET Gener., Transmiss. Distrib.*, vol. 10, no. 15, pp. 3764–3776, 2016.

[17] G. Magdy, E. A. Mohamed, G. Shabib, A. A. Elbaset, and Y. Mitani, "SMES based a new PID controller for frequency stability of a real hybrid power system considering high wind power penetration," *IET Renewable Power Gener.*, vol. 12, no. 11, pp. 1304–1313, 2018.

[18] G. Magdy, A. Bakeer, G. Shabib, A. A. Elbaset, and Y. Mitani, "Decentralized model predictive control strategy of a realistic multi power system automatic generation control," in *Proc. IEEE 19th Int. Middle East Power Syst. Conf.*, 2017, pp. 190–196.

[19] G. Magdy, G. Shabib, A. A. Elbaset, and Y. Mitani, "A novel coordination scheme of virtual inertia control and digital protection for microgrid dynamic security considering high renewable energy penetration," *IET Renewable Power Gener.*, vol. 13, no. 3, pp. 462–474, 2019.

[20] S. Sondhi and Y. V. Hote, "Fractional order PID controller for load frequency control," *Energy Convers. Manage.*, vol. 85, pp. 343–353, 2014.

[21] A. X. R. Irudayaraj et al., "A matignon's theorem based stability analysis of hybrid power system for automatic load frequency control using atom search optimized FOPID controller," *IEEE Access*, vol. 8, pp. 168751–168772, 2020.

[22] M.-R. Chen, G.-Q. Zeng, Y.-X. Dai, K.-D. Lu, and D.-Q. Bi, "Fractional-order model predictive frequency control of an islanded microgrid," *Energies*, vol. 12, no. 1, 2018, Art. no. 84.

[23] K. Mentesidi, R. Garde, M. Aguado, and E. Rikos, "Implementation of a fuzzy logic controller for virtual inertia emulation," in *Proc. IEEE Int. Symp. Smart Electric Distrib. Syst. Technol.*, 2015, pp. 606–611.

[24] T. Kerdphol, M. Watanabe, K. Hongesombut, and Y. Mitani, "Self-adaptive virtual inertia control-based fuzzy logic to improve frequency stability of microgrid with high renewable penetration," *IEEE Access*, vol. 7, pp. 76071–76083, 2019.

[25] V. Skiparev, J. Belikov, and E. Petlenkov, "Reinforcement learning based approach for virtual inertia control in microgrids with renewable energy sources," in *Proc. IEEE PES Innov. Smart Grid Technol. Eur.*, The Hague, Netherlands, 2020, pp. 1020–1024.

[26] V. Skiparev, J. Belikov, and E. Petlenkov, "MIMO reinforcement learning based approach for frequency support in microgrids with high renewable energy penetration," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, 2021, pp. 1–5.

[27] T. Kerdphol, F. Rahman, Y. Mitani, K. Hongesombut, and S. Küfeoğlu, "Virtual inertia control-based model predictive control for microgrid frequency stabilization considering high renewable energy integration," *Sustainability*, vol. 9, no. 5, 2017, Art. no. 773.

[28] N. Sockeel, J. Gafford, B. Papari, and M. Mazzola, "Virtual inertia emulator-based model predictive control for grid frequency regulation considering high penetration of inverter-based energy storage system," *IEEE Trans. Sustain. Energy*, vol. 11, no. 4, pp. 2932–2939, Oct. 2020.

[29] A. Tepljakov, E. Petlenkov, and J. Belikov, "FOMCON: Fractional order modeling and control toolbox for MATLAB," in *Proc. 18th Int. Conf. Mixed Des. Integr. Circuits Syst.*, Gliwice, Poland, 2011, pp. 684–689.

[30] A. Tepljakov, E. Petlenkov, and J. Belikov, "FOMCON toolbox for modeling, design and implementation of fractional-order control systems," in *Proc. Appl. Control*, 2019, pp. 211–236.

[31] A. Tepljakov et al., "Towards industrialization of FOPID controllers: A survey on milestones of fractional-order control and pathways for future developments," *IEEE Access*, vol. 9, pp. 21016–21042, 2021.

[32] S. Zhang et al., "Research on excitation current control system of the 50 kA superconducting transformer," *IEEE Trans. Appl. Supercond.*, vol. 31, no. 8, pp. 1–4, Nov. 2021, Art no 5501104.

[33] P. Vasant, *Meta-Heuristics Optimization Algorithms in Engineering Business Economics and Finance*. Hershey, PA, USA: IGI Global, 2013.

[34] S. D. Roy and S. Debbarma, "Frequency control in microgrid communities using neural networks," in *Proc. 51st North Amer. Power Symp.*, Wichita, KS, USA, 2019, pp. 1–6.

[35] T. Kerdphol, F. S. Rahman, M. Watanabe, Y. Mitani, D. Turschner, and H.-P. Beck, "Enhanced virtual inertia control based on derivative technique to emulate simultaneous inertia and damping properties for microgrid frequency regulation," *IEEE Access*, vol. 7, pp. 14422–14433, 2019.

[36] P. Kundur, *Power System Stability and Control*. New York, NY, USA: McGraw-Hill, 1994.

[37] H. Bevrani, B. François, and T. Ise, *Microgrid Dynamics and Control*, 1st ed. Hoboken, NJ, USA: Wiley, 2017.

[38] L. Sun and X. Zhao, "Modelling and analysis of frequency-responsive wind turbine involved in power system ultra-low frequency oscillation," *IEEE Trans. Sustain. Energy*, vol. 13, no. 2, pp. 844–855, Apr. 2021.

[39] Y. Wang, J. Meng, X. Zhang, and L. Xu, "Control of PMSG-based wind turbines for system inertial response and power oscillation damping," *IEEE Trans. Sustain. Energy*, vol. 6, no. 2, pp. 565–574, Apr. 2015.

[40] T. Lillicrap et al., "Continuous control with deep reinforcement learning," in *Proc. Int. Conf. Learn. Representations*, 2016.

[41] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, 2nd ed. Cambridge, MA, USA: The MIT Press, 2018.

[42] I. Podlubny, "Fractional-order systems and PI λ Dμ-controllers," *IEEE Trans. Autom. Control*, vol. 44, no. 1, pp. 208–214, Jan. 1999.

[43] L. Liu, F. Pan, and D. Xue, "Variable-order fuzzy fractional PID controller," *ISA Trans.*, vol. 55, pp. 227–233, 2015.

[44] P. Ostalczyk and P. Duch, "Variable, fractional-order PID controller synthesis novelty method," in *Control Based on PID Framework*, W. Wang, Ed. Rijeka, Croatia: IntechOpen, 2021, ch. 1, pp. 1–16. [Online]. Available: https://doi.org/10.5772/intechopen.95232

[45] M. Kusy and R. Zajdel, "Application of reinforcement learning algorithms for the adaptive computation of the smoothing parameter for probabilistic neural network," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 9, pp. 2163–2175, Sep. 2015.

[46] D. Silver, G. Lever, N. Heess, T. Degris, D. Wierstra, and M. Riedmiller, "Deterministic policy gradient algorithms," in *Proc. Int. Conf. Mach. Learn.*, Beijing, China, 2014, pp. 387–395.

[47] V. Skiparev, J. Belikov, E. Petlenkov, and Y. Levron, "Reinforcement learning based MIMO controller for virtual inertia control in isolated microgrids," in *Proc. IEEE PES Innov. Smart Grid Technol. Conf. Eur.*, Novi Sad, Serbia, 2022, pp. 1–5.

Appendix 7

III

K. Nosrati, V. Skiparev, A. Tepljakov, E. Petlenkov, J. Belikov, and Y. Levron. Constrained intelligent frequency control in an AC microgrid: An online reinforcement learning based PID tuning approach. In *2023 IEEE Power and Energy Society General Meeting*, pages 1–5, 2023

Constrained Intelligent Frequency Control in an AC Microgrid: An Online Reinforcement Learning Based PID Tuning Approach

K. Nosrati, A. Tepljakov, E. Petlenkov
Department of Computer Systems
Tallinn University of Technology
Tallinn, Estonia
{komeil.nosrati, aleksei.tepljakov,
eduard.petlenkov}@taltech.ee

V. Skiparev, J. Belikov
Department of Software Science,
Tallinn University of Technology,
Tallinn, Estonia
{vjatseslav.skiparev,
juri.belikov}@taltech.ee

Y. Levron
Faculty of Electrical & Computer Engineering,
Technion—Israel Institute of Technology,
Haifa, 3200003, Israel
yoashl@ee.technion.ac.il

Abstract—**Variable output power in isolated microgrids (MGs) threatens frequency stability and may even degrade power quality. In response, intelligent control methods have been developed and applied to frequency deviation control systems with excellent results. Nevertheless, a potential problem is that the application of such advanced techniques with a large search space is not enough to deal with highly dynamic environment and real-time operations of MGs. In this light, the present study introduces a flexible artificial neural network (ANN)-based frequency deviation control solution in a constrained structure that operates as follows. First, the stable controller parameter space of the PID-based AC microgrid is derived by using the stability boundary locus method. Then, the controller parameters are tuned and updated online by searching for an optimal combination of the coefficients with consideration of output variations sensed by a constrained ANN in the derived reduced parameter space. To accomplish this step, a reinforcement learning technique is applied to train the ANN-based tuners. The performance of the proposed technique has been verified under a given scenario to demonstrate how the reduced parameter space should facilitate the optimization procedure.**

Index Terms—**AC microgrid, Load frequency control, Constrained neural networks, Reinforcement learning**

I. INTRODUCTION

Technical issues in renewable energy sources (RESs) require many standards, including frequency regulation and proper control system design to achieve the desired performance. These standards are integrated into the microgrid (MG) concept [1]. Reliability enhancement, improvement in environmental issues, and economic interests are three major reasons for the emergence of MGs [2]. Small generating units, such as diesel engine generators (DEGs), photovoltaic (PV) systems, wind turbine generators (WTGs), and fuel cells (FCs), installed at the client site and connected into the power grid—referred as distributed generation—are the primary power sources in MGs [3]. Using energy storage systems (ESSs) including

battery ESS (BESS) and flywheel ESS (FESS) as backup devices enhances the stability of the MG systems, which can be degraded by the light inertia of the units [4].

Two key system parameters, i.e., voltage and frequency, must be regulated under proper control techniques due to intermittent power generation, erratic changes in load demand, and low inertia [5]. To preserve stability and improve performance, three main frequency deviation control (FDC) structures were proposed: central, single agent, and decentralized control [2]. While in the first approach, a central control takes over the FDC task according to collected data through the system, the two other methods try to use one or several controllable and locally distributed generation units for the system parameters and control purposes [6]. Modern MG systems should be able to handle complex multi-objective regulation optimization problems, which instigate the system to use intelligent FDC units and tuning methods associated with adaptable intelligent algorithms at their core [7].

To improve the robustness of the FDC problem, intelligent PID-type controllers were proposed in isolated MGs [3], [8], among many other controllers such as model predictive control [9], adaptive control [10], fuzzy logic control [11], H_∞ control [12] and conventional PID control [13], [14] approaches applied to the distributed generations. Table I summarizes some recent FDC works with intelligent optimization approaches, including the genetic algorithm (GA) [15], biogeography-based optimization (BBO) [16], particle swarm optimization (PSO) [17], reinforcement learning (RL) [18], harmony search algorithm (HSA) [8], and black hole algorithm (BHA) [3] for optimal tuning of PID-type parameters towards more efficient MG operations. To achieve better control results, an approach based on artificial neural networks (ANNs) is also applied as a tuning strategy due to its learning ability and adaptability with parameter variation [19], [20].

All these works attempted to achieve a higher degree of frequency deviation control by adjusting parameters with an offline procedure, resulting in non-flexible approaches that might fail in "dynamic adaptation" associated with a complex

Paper	Method	Controller	MG components
[15]	GA	PI/PID	WTG, STP, AE, DEG, ESS
[16]	BBO	PID	WTG, FC, PV, ESSs, AE
[2], [17]	PSO	(FO)PID	WTG, FC, PV, ESSs, DEG
[18]	RL+Fuzzy	PID	WTG, PV, FC, DEG, AE, ESSs
[3]	BHA+Fuzzy	(FO)PID	WTG, DEG, EV, BESS, FESS
[8]	HSA+Fuzzy	PI	WTG, DEG, EV
[19]	ANN	PI	WTG, PV, FC, DEG, ESSs
[20]	PSO+ANN	PID	WTG, DEG, EV
[21]	RL+ANN	PI	WTG, PV, FC, BESS

*EV: Electric Vehicle, STP: Solar Thermal Power, AE: Aqua Electrolyser
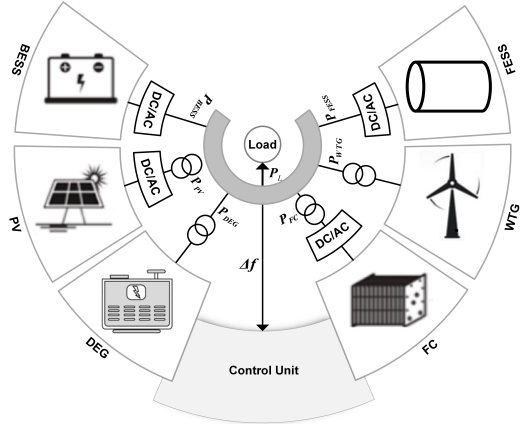
microgrid environment. This issue instigated us to propose a flexible and online-tuned PID-based FDC strategy using ANN trained by the stochastic RL (SRL) technique [21]. However, the large search space of this algorithm requires powerful computational units to backup dynamic microgrid environments. In this light, we propose here a new flexible PID-based FDC strategy, used for secondary frequency control in an AC microgrid, with automatic manipulation of all the controller's features inside the stable region obtained by the stability boundary locus (SBL) method. In the proposed control strategy, the constrained ANN-based architecture tunes the controller in a reduced space where the weights in each neuron are trained by using an RL technique on the specified stable region, as opposed to the entire space. We discovered how the proposed online tuning method offers many benefits for the FDC system in an AC microgrid with the reduced parameter space, where it shall facilitate the optimization procedure for the network training as the search space becomes smaller.

II. STRUCTURE OF THE AC MICROGRID

In Fig. 1, an isolated AC microgrid system, including a standard DEG, PV panels, WTGs, a FC system, a BESS, and a FESS, is depicted associated with power electronic interfaces. Fuel blocks, a DC/AC inverter, and interconnection devices are all included in the FC system. Although the FC has a high-order characteristic, frequency investigations only require a three-order model [22]. Nominal values of rated power for the subsystems DEG, RESs, FC, and ESSs are 160, 130, 70, and 90 kW, respectively. The DEG is required to provide a certain amount of power that is designated as a spinning reserve for secondary frequency regulation. A block diagram of the simplified frequency response model for the case study is shown in Fig. 2 associated with the given parameters in Table II, where P_{Wind} and φ are the mechanical power of the WTG and solar irradiation, respectively.

In the illustrated model and to control the system frequency, the DEG and FC are regarded as manipulating units. Also, due to the variable and non-measurable nature of input powers of PV, WTG, and load (denoted as P_{PV}, P_{WTG}, and P_L, respectively), they can be considered as input disturbance signal obtained as $P_{tot,L} = P_L - P_{PV} - P_{WTG}$,

Fig. 1. General configuration of the isolated AC microgrid.
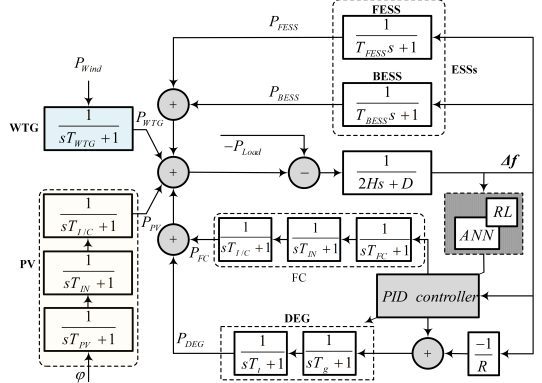
Fig. 2. Frequency block diagram of the AC microgrid system.

TABLE II
PARAMETERS OF THE AC MICROGRID

Parameter	Value	Parameter	Value
D	0.0015 pu/Hz	T_g	0.08 s
$2H$	0.1667 pu.s	T_t	0.4 s
T_{FESS}	0.1 s	$T_{I/C}$	0.004 s
T_{BESS}	0.1 s	T_{IN}	0.04 s
T_{FC}	0.26 s	R	3 Hz/pu
T_{WTG}	1.5 s	T_{PV}	1.8 s

where $P_{tot,L}$ is the generalized load power used to derive the characteristic equation in the following section.

III. STABILITY ANALYSIS OF THE AC MICROGRID

For stability analysis of the given microgrid system, the characteristic equation of the system will be derived by forming the closed loop transfer function as $T(s) = \frac{\Delta f}{P_{tot,L}}$,

and its roots will be checked regarding to the complex root boundary (CRB). To this end, a PID controller

$$G_C = K_p + \frac{K_i}{s} + K_d s, \tag{1}$$

where K_p, K_I, and K_d are proportional, integral and derivative control gains, is applied to design the FDC system for manipulating the DEG and FC units. After some mathematical analysis, the denominator of the transfer function $T(s)$ is the characteristic equation $\Delta(s)$ expressed as

$$\begin{aligned}
\Delta(s) = {} & 6.66 \times 10^{-9} s^9 + 2.0894 \times 10^{-6} s^8 \\
& + (1.25 \times 10^{-6} K_d + 1.1731 \times 10^{-4}) s^7 \\
& + (0.0013 K_d + 1.25 \times 10^{-6} K_p + 0.0029) s^6 \\
& + (0.0498 K_d + 1.25 \times 10^{-6} K_i + 0.0013 K_p + 0.042) s^5 \\
& + (0.66 K_d + 0.0013 K_i + 0.0498 K_p + 0.3788) s^4 \\
& + (3.552 K_d + 0.0498 K_i + 0.66 K_p + 2.0188) s^3 \\
& + (6 K_d + 0.6612 K_i + 3.552 K_p + 5.3045) s^2 \\
& + (3.552 K_i + 6 K_p + 5.0045) s + 6 K_i.
\end{aligned} \tag{2}$$

In order to perform the required stability analyses, the characteristic equation (2) will be rewritten by replacing $s = j\omega$. As a result, we can arrange the obtained equation in a complex polynomial form as

$$\Delta(\omega, K_{p,i,d}) = \begin{bmatrix} 1 \\ j \end{bmatrix}^{\mathsf{T}} \left(\begin{bmatrix} A_1(w, K_d) & A_2(w, K_d) \\ B_1(w, K_d) & B_2(w, K_d) \end{bmatrix} \begin{bmatrix} K_p \\ K_i \end{bmatrix} + \begin{bmatrix} A_3(w, K_d) \\ B_3(w, K_d) \end{bmatrix} \right),$$

where

$$A_1(w, K_d) = -3.552\omega^2 + 0.05\omega^4 - 1.25 \times 10^{-6}\omega^6,$$
$$A_2(w, K_d) = 6 - 0.66\omega^2 + 0.0013\omega^4,$$
$$\begin{aligned}A_3(w, K_d) = {} & -(5.3045 + 6K_d)\omega^2 + (0.66K_d + 0.3788)\omega^4 \\ & - (0.0029 + 0.0013K_d)\omega^6 + 2.0894 \times 10^{-6}\omega^8,\end{aligned}$$
$$B_1(w, K_d) = 6\omega - 0.6612\omega^3 + 0.0013\omega^5,$$
$$B_2(w, K_d) = 3.552\omega - 0.045\omega^3 + 1.25 \times 10^{-6}\omega^5,$$
$$\begin{aligned}B_3(w, K_d) = {} & 5.0045\omega - (2.0188 + 3.552K_d)\omega^3 + (0.042 \\ & + 0.05K_d)\omega^5 - (1.25 \times 10^{-6}K_d + 1.2 \times 10^{-4}) \\ & \times \omega^7 + 6.7 \times 10^{-9}\omega^9.\end{aligned}$$

By setting the imaginary and real parts of the last equation equal to zero, we can derive two parameters of the controller regarding the derivative control gain K_d and frequency ω. In this case, by considering K_d as a fixed parameter, we have

$$\begin{aligned}
K_p &= \frac{A_2(w, K_d)B_3(w, K_d) - A_3(w, K_d)B_2(w, K_d)}{A_1(w, K_d)B_2(w, K_d) - A_2(w, K_d)B_1(w, K_d)}, \\
K_i &= \frac{A_3(w, K_d)B_1(w, K_d) - A_1(w, K_d)B_3(w, K_d)}{A_1(w, K_d)B_2(w, K_d) - A_2(w, K_d)B_1(w, K_d)}.
\end{aligned} \tag{3}$$

Now, according to the defined boundaries as

$$\Delta(j\omega, K_{p,i,d}) = 0 \Rightarrow \begin{cases} \omega = 0, & K_i = 0 \\ 0 < \omega < \infty, & K_{p,i} \in CRB \\ \omega = \infty, & K_{p,i} \notin CRB \end{cases} \tag{4}$$
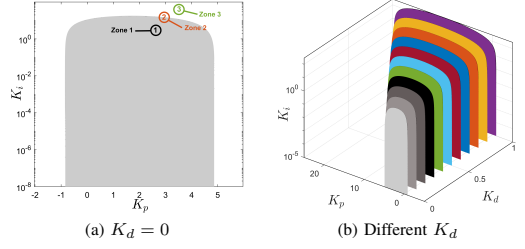

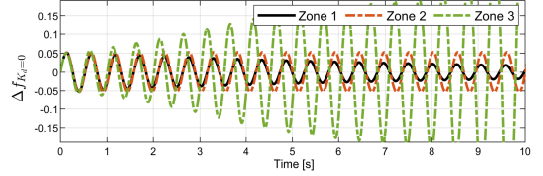Fig. 3. SBL of PID controller for the isolated AC MG.

Fig. 4. Frequency deviation in different zones of Fig. 3a.

the SBL curves can be drawn in the $K_p - K_i$ plane with respect to K_d as shown in Fig. 3. Also, the frequency response of the given AC microgrid, for different three zones (outside, on and within the SBL curve) and when $K_d = 0$, is depicted in Fig. 4. For these three zones 3, 2 and 1, the microgrid is unstable, marginally stable, and stable, respectively.

IV. CONSTRAINED NN-PID BASED TUNING ALGORITHM

In this section, we will introduce the architecture of ANN, which is designed with constraints defined by stability region to not only guarantee the stability of the PID controller in the microgrid but also reduce the parameter space to facilitate the optimization procedure for the network training. This strategy will be limited by the range of ANN weights, constants for the activation function, and tuner output values.

A. SRL-based Training for Constrained NN-PID Controller

To perform the training of the constrained NN-PID controller, we apply the basic SRL algorithm proposed in [21] and use it to train a designed agent to provide optimal tuning of the PI part of the PID controller. The training is based on reward feedback organized by the reward system (see Eq. 5), where the optimal weights can be obtained via the selection of actions with the highest possible reward. In this approach, we perform actions a_t (i.e., $K_p(t)$ and $K_i(t)$ for a fixed K_d) and check the received maximum average reward $r_{\text{avg,max}}$ of the RL agent. By using the regulation system

$$r_t = \begin{cases} \frac{1}{0.5 + 100|\Delta f|}, & \text{if } 100|\Delta f| < 0.05 \\ -200|\Delta f|, & \text{if } 100|\Delta f| > 0.05, \end{cases} \tag{5}$$

where the main criterion on which rewarding based is the magnitude of control error Δf, the agent receives a positive reward $+r$ if $\Delta f < 0.05$ Hz; otherwise, it takes a negative

reward or punishment $-r$. According to the applied SRL optimization approach, agent will be defined as $\mu(s \mid \theta_{\mu_i})$ with initialized random weights $\sum_{j=1}^{M} \theta_{\mu_j}$, where θ is array of non-negative random values that initializes in each episode M. Then, the reward function $r_t(a_t, s_t)$, where $a_t = K_{PI}(t)$ and $s_t = \Delta f_t$, will be observed at each action and state. Finally, the weights θ_μ of agent with best average reward will be selected by

$$\theta_\mu = \begin{cases} \theta_{\mu_{M+1}} = \theta_{\mu_M}, & \text{if } r_{\text{avg}} > r_{\text{avg,max}}, \\ \theta_{\mu_M} = \theta_{\mu_M}, & \text{if } r_{\text{avg}} < r_{\text{avg,max}}, \end{cases} \quad (6)$$

where $r_{\text{avg,max}} = \max\left(\frac{1}{N}\sum_t r_t(s_t, a_t)\right)$. The pseudo-code of this approach is summarized in Algorithm 1.

Algorithm 1 SRL optimization approach.

1: Initialize neural network of agent as $a_t = \mu(s_t \mid \theta_\mu)$
2: **for** $j = 1$ to M **do**
3:     Initialize random weights of agent θ_{μ_M} M times
4:     Execute action a_t, observe reward r_t at each step t in common state s_t
5:     **if** Average reward $r_{\text{avg}} > r_{\text{avg,max}}$ **then**
6:         Save weights of agent θ_μ
7:     **end if**
8: **end for**

B. Activation Function

For the proposed constrained ANN-based tuner adopted from [21] and constrained to the stability validation test, we apply the activation function

$$y(t) = \frac{m}{1 + e^{-0.5|x(t)|}} - \frac{m}{2}, \quad (7)$$

where constant m is the limiting feature for the range of every ANN output. When $K_d = 0$, the output range of two other parameters are as $K_p \in [0, \ldots, 5]$ and for $K_i \in [0, \ldots, 16]$. As seen in Fig. 5, while Fig. 5a shows ANN-based tuning of the controller without stability analysis, Fig. 5b presents the validated design of an ANN that "passes" the stability test, where as opposed to the former, the automatic manipulation of all parameters is limited to the reduced stable space. Since the search space becomes smaller, this can facilitate the optimization procedure for the network training.
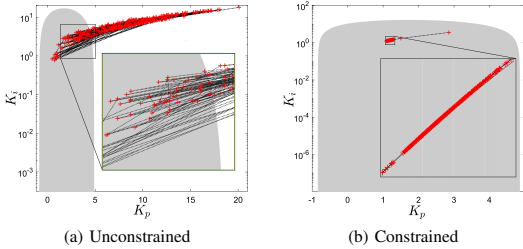

(a) Unconstrained   (b) Constrained

Fig. 5. Validation of the proposed tuner by (un)constrained ANN.

V. NUMERICAL RESULTS

MATLAB/Simulink was used to build the proposed strategy and test the performance of the constrained NN-PID controller using the given FDC scheme. Figure 6 shows the power of RESs with disconnection after 140 seconds and connection at 180 seconds for WTG and PV, respectively, load demand with disconnection and connection at 170 and 190 seconds, respectively. Figure 7 and Table III show comparative results and demonstrate how effectively the proposed algorithm can solve the FDC problem and shows more smooth deviation as compared to the usual PID controller with fixed parameters $K_p = 3.712$ and $K_i = 1.391$ for the case when K_d is zero. These optimal parameters have been obtained by heuristic search to provide a comparison with the proposed method. As can be seen, in both transient and steady state response, the constrained NN-PID controller contributes more effective results. By increasing K_d, the performance is rather comparable which can be negligible since the PI controller is preferred in the industry due to the derivative component typically contributing to amplifying noise. Figure. 8 shows the dynamics of controller parameters in the online tuning strategy of the proposed method in the given scenario.
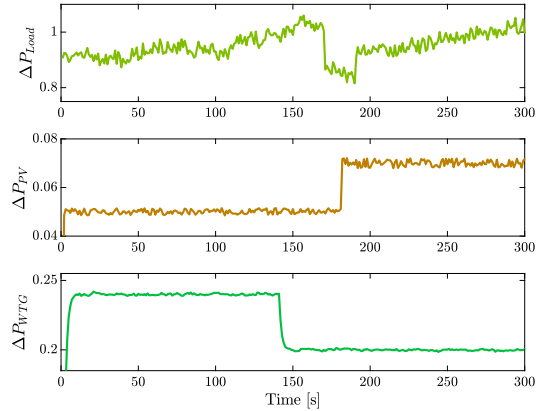
Fig. 6. Microgrid profile of load, wind and solar power.

TABLE III
NUMERICAL RESULTS OF COMPARED PID CONTROLLERS

K_d	Controller	IAE	RMSE	MAE
1.0	NN-PID	**0.9061**	0.0042	0.0031
	PID	0.9066	**0.0038**	0.0031
0.5	NN-PID	**1.2960**	0.0053	**0.0044**
	PID	1.3853	0.0058	0.0047
0.0	NN-PI	**2.757**	**0.0096**	**0.0078**
	PI	3.9297	0.0167	0.0133
-	Without	80.5724	0.2744	0.2722

*IAE: Integral Absolute Error, MAE: Mean AE,
RMSE: Root Mean Square Error
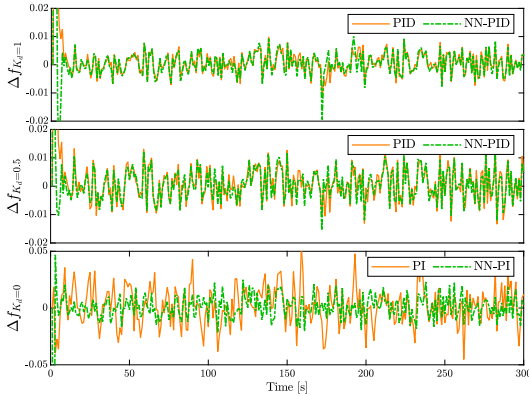
Fig. 7. Numerical comparison of the tested algorithms.
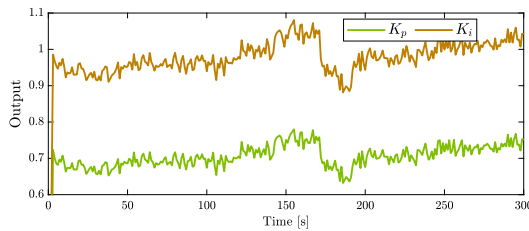
Fig. 8. Online tuning of the parameters with $K_d = 0$.

CONCLUSION

In this work, we propose a flexible and online-tuned PID-based frequency deviation control strategy using an ANN trained by the SRL technique. To derive this dynamic adaptable controller for the well-supported dynamic AC microgrid environments, a constrained ANN-based tuning architecture was introduced, in which the algorithm tries to meaningfully tune the controller in a reduced stable space, as opposed to the entire space. The simulation results verified how the proposed method offers many benefits for the frequency deviation control of the microgrid, with automatic manipulation of all the controller's features inside the reduced parameter space, where it facilitates the optimization procedure for the network training. Future work could apply non-integer controllers associated with communication delays and more accurate training space to the proposed method to increase the tuning freedom together with the microgrid stability that leads to fulfilling intricate control performance requirements.

REFERENCES

[1] R. H. Lasseter, J. H. Eto, B. Schenkman, J. Stevens, H. Vollkommer, D. Klapp, E. Linton, H. Hurtado, and J. Roy, "CERTS microgrid laboratory test bed," IEEE Trans. Power Del., vol. 26, pp. 325–332, Jan. 2011.

[2] H. Bevrani, F. Habibi, P. Babahajyani, M. Watanabe, and Y. Mitani, "Intelligent frequency control in an AC microgrid: Online PSO-based fuzzy tuning approach," IEEE Trans. Smart Grid, vol. 3, no. 4, pp. 1935–1944, Dec. 2012.

[3] M. Khooban, T. Niknam, M. Shasadeghi, T. Dragicevic, and F. Blaabjerg, "Load frequency control in microgrids based on a stochastic noninteger controller," IEEE Trans. Sustain. Energy, vol. 9, no. 2, pp. 853–861, Apr. 2018.

[4] S. Jain and Y. V. Hote, "Design of Improved Nonlinear Active Disturbance Rejection Controller for Hybrid Microgrid With Communication Delay," IEEE Trans. Sustain. Energy, vol. 13, pp. 1101–1111, Feb. 2022.

[5] I. Pan and S. Das, "Kriging based surrogate modeling for fractional order control of microgrids," IEEE Trans. Smart Grid, vol. 6, no. 1, pp. 36–44, Jan. 2015.

[6] M. Datta, T. Senjyu, A. Yona, T. Funabashi, and C. H. Kim, "A coordinated control method for leveling PV output power fluctuations of PV-diesel hybrid systems connected to isolated power utility," IEEE Trans. Energy Convers., vol. 24, pp. 153–162, Mar. 2009.

[7] H. Bevrani and T. Hiyama, Intelligent Automatic Generation Control. New York: CRC, 2011.

[8] M. H. Khooban, T. Niknam, F. Blaabjerg, and T. Dragicevi, "A new load frequency control strategy for micro-grids with considering electrical vehicles," Elect. Power Syst. Res., vol. 143, pp. 585–598, Feb. 2017.

[9] A. K. Verma, H. B. Gooi, A. Ukil, N. R. Tummuru, and S. K. Kollimalla, "Microgrid frequency stabilization using model predictive controller," in Proc. IEEE PES Transm. Distrib. Conf. Expo. Latin America (PES T D-LA), Sep. 2016, pp. 1–6.

[10] S. Zheng, X. Tang, B. Song, S. Lu, and B. Ye, "Stable adaptive PI control for permanent magnet synchronous motor drive based on improved JITL technique," ISA Trans., vol. 52, no. 4, pp. 539–549, July 2013.

[11] D. Q. Oliveira, A. C. Souza, M. V. Santos, A. B. Almeida, B. I. L. Lopes and O. R. Saavedra, "A fuzzy-based approach for microgrids islanded operation," Electr. Power Syst. Res., vol. 149, pp. 178–189, Aug. 2017.

[12] H. Bevrani, M. R. Feizi and S. Ataee, "Robust frequency control in an islanded microgrid: H_∞ and μ-synthesis approaches," IEEE Trans. Smart Grid, vol. 7, pp. 706–717, Aug. 2016.

[13] M. Nayeripour, M. Hoseintabar and T. Niknam, "Frequency deviation control by coordination control of FC and double-layer capacitor in an autonomous hybrid renewable energy power generation system," Renew. Energ., vol. 36, pp. 1741–1746, June 2011.

[14] K. Nosrati, H. R. Mansouri and H. Saboori, "Fractional-order PID controller design of frequency deviation in a hybrid renewable energy generation and storage system," CIRED - Open Access Proc. J., vol. 1, pp. 1148–1152, Oct. 2017.

[15] C. D. Das, N. Sinha and A. K. Roy, "GA based frequency controller for solar thermal-diesel-wind hybrid energy generation/energy storage system," Int. J. Elec. Power, vol. 43, pp. 262–279, Dec. 2012.

[16] H. Kumar and S. Ushakumari, "Biogeography based Tuning of PID Controllers for Load Frequency Control in Microgrid," in Proc. Int. Conf. Circuit, Power and Comput. Tech. (ICCPCT), Mar. 2014, pp. 797–802.

[17] I. Pan and S. Das, "Fractional order AGC for distributed energy resources using robust optimization," IEEE Trans. Smart Grid, vol. 7, pp. 2175–2186, Aug. 2016.

[18] M. Esmaeili, H. Shayeghi, H. M. Nejad, and A. Younesi, "Reinforcement learning based PID controller design for LFC in a microgrid," Int. J. Comput. Math. Electr. Electron. Eng., vol. 36, pp. 1287–1297, July 2017.

[19] H. Bevrani, F. Habibi and S. Shokoohi, "ANN-based self-tuning frequency control design for an isolated microgrid," Meta-heuristics optimization algorithms in engineering business economics and finance, pp. 357, 2012.

[20] H. Kumar and S. Ushakumari, "Frequency control in microgrid communities using neural networks," in 2019 North American Power Symposium (NAPS). IEEE, Feb. 2019, pp. 1–6.

[21] K. Nosrati and V. Skiparev and A. Teplajkov and E. Petlenkov and Y. Levron and J. Belikov, "Coordinated PI-based frequency deviation control of isolated hybrid microgrid: An online multi-agent tuning approach via reinforcement learning," 2022 IEEE PES Innovative smart grid technologies (ISGT-Europe), Oct. 2022.

[22] S. Obara, "Analysis of a fuel cell micro-grid with a small-scale wind turbine generator," Int. J. Hydrog. Energy, vol. 32, pp. 323–336, Mar. 2007.

Appendix 8

II

V. Skiparev, K. Nosrati, J. Belikov, A. Tepljakov, and E. Petlenkov. An enhanced NN-based load frequency control design of MGs: A fractional order modeling method. In *2023 IEEE International Conference on Compatibility, Power Electronics and Power Engineering (CPE-POWERENG)*, Tallinn, Estonia, 2023

An Enhanced NN-based Load Frequency Control Design of MGs: A Fractional order Modeling Method

Vjatseslav Skiparev
Department of Software Science
Tallinn University of Technology
Tallinn, Estonia
vjatseslav.skiparev@taltech.ee

Komeil Nosrati
Department of Computer Systems
Tallinn University of Technology
Tallinn, Estonia
komeil.nosrati@taltech.ee

Juri Belikov
Department of Software Science
Tallinn University of Technology
Tallinn, Estonia
juri.belikov@taltech.ee

Aleksei Tepljakov
Department of Computer Systems
Tallinn University of Technology
Tallinn, Estonia
aleksei.tepljakov@taltech.ee

Eduard Petlenkov
Department of Computer Systems
Tallinn University of Technology
Tallinn, Estonia
eduard.petlenkov@taltech.ee

Abstract—Developing accurate mathematical models for microgrid (MG) components is the initial step before implementing various load frequency control (LFC) strategies and analysis. In this regard, different high-order models associated with different nonlinearities have been included to increase the modeling accuracy resulted in a performance improvement in the LFC techniques. Nevertheless, these high-order nonlinear models pose some potential problems such as obstacles in the analytical description of the system and control problem along with its high computational complexity. In this light, the fractional order based models are deployed to effectively balance the model accuracy and analytical complexity. First, two fractional order components (energy storage system and fuel cell) are arranged in a controlled coordinated strategy to enhance the frequency stability. Then, two artificial neural network (ANN) controllers are deployed for each components in a multi-agent framework. To accomplish this step, a multi-agent stochastic reinforcement learning optimization is applied to train the two controllers. Test results on an isolated MG with fractional components validate the efficacy of the coordinated LFC strategy.

Index Terms—Load frequency control, multi-agent, neural networks, reinforcement learning, fractional order

I. INTRODUCTION

A microgrid (MG) system is a localized power grid that can operate autonomously or connected to a larger electrical grid [1]. In MGs, wind turbine generators (WTGs) and photovoltaic (PV) systems are promising renewable technologies for generating electricity [2]. To ensure a stable power supply and mitigate fluctuations, MGs integrate energy storage systems (ESSs) and additional generation components such as fuel cells (FCs) [3]. While FC systems perform well under steady-state conditions, they have slow dynamics and a time delay, so they are often used together with ESSs to achieve the desired power output [4]. Due to the limited control capabilities of FCs, double-layered capacitors (DLCs) are used to improve the transient response of the system [5]. By employing a combination of these subsystems, MGs can provide reliable and sustainable power to local communities [4].

A hybrid MG is a power system that combines two or more power generation sources to improve its overall performance [6]. To ensure efficient power delivery to the load sides, it is important to coordinate the frequency deviation of the system's components [7]. Various methods for load frequency control (LFC) system have been proposed [8], including droop control [9], model predictive control [10], fuzzy logic control [11], and H_∞ control [12]. A new co-ordinated (fractional order) PID (FOPID) controller, which has recently received the attention of researchers due to its ability to be fine-tuned using extra adjustable parameters [13], has been developed for an isolated MG [14]–[16], which offers some enhancements over previous methods. However, by optimizing the parameters of the PID-type controllers using modern approaches such as genetic algorithm , particle swarm optimization, and reinforcement learning (RL), the stability of hybrid MG operation can be significantly improved [17].

To achieve better control results, accurate mathematical models together with modern adaptable control approaches can be exploited for LFC systems. From one side, artificial neural networks (ANN) have been assigned as a well-suited candidate for LFC purposes in a highly dynamic MG environment due to its adaptability and learning ability [17]–[19]. On the other

hand, different high-order models associated with different nonlinearities have been included to increase the modeling accuracy resulted in a performance improvement in the LFC systems [20]. Notwithstanding this, these high-order nonlinear models pose some potential problems during analysis, design, and verification process of the control system. One disadvantage is that such nonlinear models contribute major obstacles in the analytical description of the system and control problem due to limitations of nonlinear mathematical tools. In addition, such high-order models lead to an increase in computational burdens and required processing time. The aforementioned problems may be crucial in practical load frequency control of complex MG environment under its intermittent dynamics and variable operating points.

In this light, the present work exploit the fractional order modeling techniques to effectively address a trade-off between model accuracy and analytical complexity. Under this balanced strategy, two fractional order components (DLC and FC) are arranged in a coordinated strategy to enhance the frequency stability. Then, two ANN controllers are deployed for each components in a multi-agent framework where a common stochastic RL (SRL) optimization method is applied to train each controllers. Our strategy will facilitate model-based control for unlocking the full potential of the LFC system. Based on the test results, we discovered that it leads to a superior performance of the LFC, including improvements in frequency stability and robustness.

II. CONFIGURATION OF MG SYSTEM

A. General Structure

Figure 1 presents the proposed isolated MG, which encompasses power generation subsystems consisting of a WTG, PV panels, FC system, and a DLC. These subsystems are interconnected in parallel to a shared AC bus that provides isolated loads. To examine this system, high-order mathematical models with nonlinear dynamics are required for each subsystem. However, in the majority of studies, first-order transfer functions are used to analyze all components, and the system simulations relied on a simplified linear models. Although deploying nonlinear modeling techniques can greatly increase the accuracy and extracting the full system potential, they come at a high computational complexity along with challenges in the LFC system analysis during the design process [20]. In this regard, we substitute the FC and DLC systems by their fractional-order models which eliminate the disadvantage of the low accuracy while at the same time reducing the model complexity.

B. WTG Model

The amount of power produced by wind turbines is directly influenced by the speed of the wind. The mechanical power generated by the turbine is determined by $P_{WTG} = 0.5\rho A v^3 C_p(\lambda, \theta)$, which takes into account the density of air ρ, the swept area of the blades A, and the wind velocity v. The rotor efficiency, or C_P, is also an important factor and is determined by $C_p(\lambda, \theta) = 0.5(116\hat{\lambda}^{-1} - 0.4\theta - 5)e^{-21\hat{\lambda}^{-1}} +$
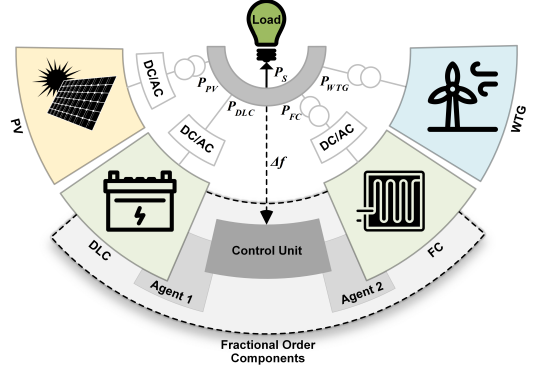

Fig. 1. General structure of the isolated MG.

0.0068λ with $\hat{\lambda}^{-1} = (\lambda + 0.08\theta) - 0.035(\theta^3 + 1)$, which is a function of the tip speed ratio λ and pitch angle θ. The maximum value of $C_{P,\max}$ can be obtained for a specific blade direction and when the tip speed ratio is at its optimal level. The rotor speed can be adjusted to keep the tip speed ratio at this level, allowing for maximum energy extraction from the wind (see Fig. 2a). As shown in Fig. 2b, the power output P_{WTG} remains constant by using a pitch angle control system to prevent excessive rotor speed and protect the equipment. The power output reaches its maximum value when the wind speed is between 14 and 25 m/s and is zero when the wind speed is less than 4 m/s. However, if the wind speed exceeds 25 m/s, the system is shut down to prevent damage to the turbine. The frequency domain's mathematical representation of a WTG's dynamics can be expressed as

$$\frac{\Delta P_{WTG}}{\Delta P_{wind}} = \frac{1}{sT_{WTG} + 1}. \quad (1)$$

This equation shows the relationship between changes in the mechanical power of the turbine, denoted by ΔP_{wind}, and the output power of the generator, represented by ΔP_{WTG}.
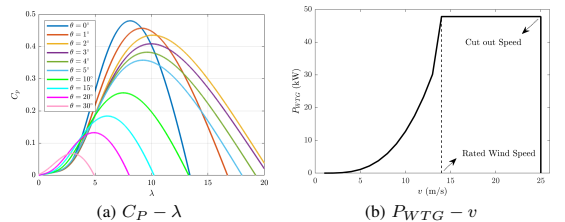

(a) $C_P - \lambda$ (b) $P_{WTG} - v$

Fig. 2. The changes in C_P and P_{WTG} with respect to λ an v.

C. PV Model

The PV system is an eco-friendly and highly versatile power source that is made up of photovoltaic panels arranged in series and parallel configurations. This system converts solar radiation into electrical energy by utilizing the temperature of

the solar cells and the array's surface area. The amount of electrical energy produced by the PV system, represented as P_{PV}, can be calculated based on the conversion efficiency η, the area of the PV array S, the solar irradiation φ, and the ambient temperature T_a, as expressed by equation

$$P_{PV} = \eta S \varphi (1 - 0.005(T_a + 25)). \tag{2}$$

The relationship between the variations in solar irradiation and the output power of the PV system is given as

$$\frac{\Delta P_{PV}}{\Delta \varphi} = \frac{1}{sT_{PV} + 1}, \tag{3}$$

where T_{PV} is the time constant of the PV system.

D. Fractional FC Model

The FC technology is known for its high efficiency in generating power by converting chemical energy into electrical energy using hydrogen and oxygen. However, the fuel supply sections of this device, such as pumps and valves, have slow dynamics, which can result in a slower power output. To analyze the dynamic response of this system, the first-order time delay transfer function can be introduced as

$$G_{FC} = \frac{1}{sT_{FC} + 1}. \tag{4}$$

In order to improve control design and by using fractional-order modeling technique, which is a versatile mathematical tool enabling non-integer order derivatives in model descriptions [21], a new fractional FC (FFC) was introduced in the study [22]. This model accurately captures all dynamic behaviors using proposed free fractional order elements. The advanced solid oxide FC model includes two constant phase element (CPE) components that describe the anode and cathode behavior of the FC. The transfer function of the FFC model can be represented as

$$G_{FFC} = \frac{1}{s^{\alpha_1} T_{FC} + 1} + \frac{1}{s^{\alpha_2} T_{FC} + 1}, \tag{5}$$

where $0 < \alpha_i < 1$, $i = 1, 2$ are fractional orders related to each CPE [23]. Compared to classic integer-order models, the presented FOFC model has the benefit of higher modeling accuracy, particularly for the dynamic behavior during transient operation.

E. Fractional DLC Model

Electro-chemical capacitors were first patented by General Electric Company in 1957, using a charging mechanism known as DLC and consisting of porous electrodes [24]. Today, this technology is widely used in a variety of applications, including power quality, electric vehicles, and renewable energy systems. The device is highly efficient, with a fast load frequency, and its transfer function can be approximated by the first-order lag equation, specifically

$$G_{DLC} = \frac{1}{sT_{DLC} + 1}. \tag{6}$$

This traditional modelling requires a large number of parameters in a larger frequency bands [23]. Therefore, researchers have explored fractional calculus potential for modeling electrochemical systems, such as batteries and supercapacitors, and have found that it can improve accuracy and reduce complexity compared to other methods. For example, a simplified electrochemical model based on the fractional-order model has fewer parameters than other approaches and can provide a more accurate estimation of battery states [25]. Other studies have investigated factors that affect the accuracy of fractional models, such as ambient temperatures, memory lengths, different profiles, and voltage/current drifts [26]. Overall, the fractional-order model shows great potential for improving the accuracy of equivalent circuit models while reducing their complexity, making it a promising tool for modeling electrochemical systems [27]. A straightforward model for fractional DLC (FDLC) can be derived from the behavior of porous electrodes in DLCs. This model is represented by the equation

$$G_{FDLC} = 1 + \frac{1}{s^{\beta} T_{DLC}}, \tag{7}$$

where $0 < \beta < 1$ is the derivative order related to the CPE.

Table I summarizes the deployed parameters in the MG system and given equations.

TABLE I
PARAMETERS OF THE MICROGRID

Parameter	Physical meaning	Nominal value	Unit
T_{DLC}	WTG time constant	0.1	s
T_{FC}	FC time constant	0.2	s
T_{WTG}	WTG time constant	1.5	s
T_{PV}	PV time constant	1.8	s
β	Derivative order of CPE in FDLC	0.5	–
α_1	Derivative order of first CPE in FFC	0.8	–
α_2	Derivative order of second CPE in FFC	0.8	–
M	Inertia constant	0.166	pu.Hz
D	Damping constant	0.015	pu/Hz

III. COORDINATED CONTROL STRATEGY

By utilizing a coordinated control strategy that combines FC and DLC systems, the LFC problem can be effectively resolved while also improving power quality, as depicted in Fig. 3. In this approach, the fractional order DLC and FC systems serve as backup systems and compensate for high and low frequency deviations, respectively. Also, to minimize the charging and discharging of FDLC during long-term operation, a high-pass filter (HPF) is employed. Furthermore, to maintain stable operation of the autonomous isolated MG system, effective control of the supply power is necessary since the output power of various power generation components can fluctuate under certain conditions. The control strategy in this system is based on the power balance error ΔP, which is calculated as the difference between the power supply P_{Net} and the power demand P_{Load} given as $\Delta P = P_{Net} - P_{Load}$.

As power generation varies, the frequency also fluctuates, and this frequency deviation Δf, is calculated using

$$\Delta f = \frac{\Delta P}{K}, \tag{8}$$

where K is the system frequency characteristic constant of the system. However, due to delays in the frequency characteristics, the above equation is modified to

$$\Delta f = \frac{\Delta P}{K(sT_f + 1)} = \frac{\Delta P}{sM + D}, \tag{9}$$

which takes into account the frequency characteristic time constant T_f, as well as the load damping constant D and the inertia constant M. The net power generation is computed using $P_{Net} = P_{WTG} + P_{PV} + P_{FFC} \pm P_{FDLC}$, which involves various contributing factors. For control purposes, two ANN based control are employed for two fractional subsystems associated with a common stochastic RL based training unit. This control framework and its strategy, which is integrated in control unit, will be more investigated in two following subsections.
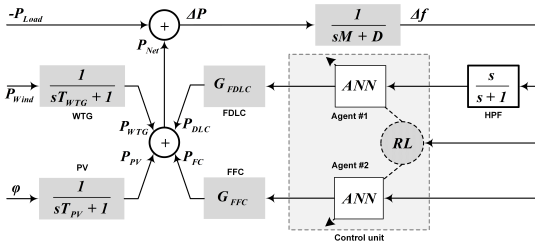

Fig. 3. Coordinated LFC strategy.

A. ANN-based Controller

The proposed NN-based controller, which is similar to PID-type controllers, is illustrated in Fig. 4. The controller takes frequency deviations as an error signal in three different forms: original Δf, differentiated $d\Delta f/dt$, and integrated $\Delta f/s$, and inputs them into neurons. In this architecture, multiple ANN weights of actor systems are defined as $w_{\mu 1,i}$ and $w_{\mu 2,i}$ for hidden and output layers, respectively. The hidden layer consists of three neurons, while the output layer has one neuron that summarizes all signals into output provided for the inputs of FDLC and FFC systems. Every artificial neuron uses an adjustable tansig activation function that provides bipolar output and is suitable for many control applications. The performance of the proposed controller depends on the magnitude of the weights in each artificial axon and the constants defined in the applied activation function as

$$y(t) = \frac{n}{1 + e^{-wx(t)}} - \frac{n}{2}, \tag{10}$$

where weight w defines sensitivity to disturbances and constant n defines limits for output signal in every neuron.
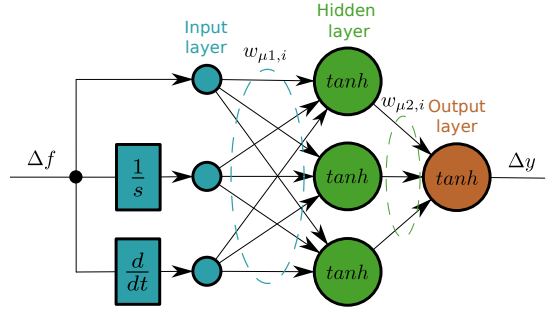

Fig. 4. Proposed ANN-based controller.

B. Stochastic Reinforcement Learning

In order to train multiple agents to control the given isolated MG, we utilize a stochastic reinforcement learning (SRL) algorithm proposed in [17]. This algorithm initializes random weights m times and evaluates each combination of weights using a reward system. The proposed reward system is based on the angular function

$$r_t = \begin{cases} 2\frac{\text{IAE}}{\sqrt{\text{IAE}^2 + t^2}}, & \text{if } \sin\theta < 0.5 \\ -10\frac{t}{\sqrt{\text{IAE}^2 + t^2}}, & \text{if } \sin\theta > 0.5 \end{cases} \tag{11}$$

that takes into account the integral absolute error (IAE) of the system's performance defined as

$$\text{IAE} = \int_0^t |\Delta f| \text{d}t. \tag{12}$$

The optimal weights are obtained through the selection of actions that yield the highest reward for each agent. This approach uses an ANN with the mentioned tansig activation function and the PID-like control approach. In this approach, we execute actions a_t at a given state s_t and assess the maximum reward $r_{\max}$ received by each agent separately. The reward is determined by the measured control error, which is represented by the magnitude of the state difference (Δf) and is calculated as the IAE signal. If the error angle (θ) is less than 45 degrees, the agent receives a positive reward ($+r$); otherwise, it receives a negative reward or punishment ($-r$). Unlike a previous approach, which had limitations for highly dynamic control applications such as interconnected MGs, the new reward system is adaptive to such systems since it is based on the performance metric IAE, which is an effective criterion for process control. However, training multiple agents can be challenging due to the interdependence of their actions. To overcome this, we calculate a common summarized reward for each agent and save the weights associated with the best achieved reward.

The proposed optimization approach for multi-agent SRL involves defining a certain number of ANN agents as $\sum_{i=1}^{l} \mu_i(s \mid w_{\mu_i})$ with the non-negative random value w and random weights $\sum_{i=1}^{l} \sum_{j=1}^{M} w_{\mu_{i,j}}$, which are modified over

a specific number of episodes (M). At each state and action, the reward function $r_t(a_{t_l}, s_t)$, where $a_{t_l} = ANN_l(t)$ and $s_t = \Delta f_t$, is observed and used to evaluate the performance of each agent. The weights associated with the highest reward for each agent are then selected, using

$$w_{\mu_l} = \begin{cases} w_{\mu_{l,M+1}} = w_{\mu_{l,M}}, & \text{if } r_{new_l} > r_{\max_l}, \\ w_{\mu_{l,M}} = w_{\mu_{l,M}}, & \text{if } r_{new_l} < r_{\max_l}, \end{cases} \quad (13)$$

that considers the maximum reward $r_{\max_l} = \max(r_{t_l}(s_t, a_{t_l}))$ received by each agent. The approach is summarized in Algorithm 1, which outlines the steps involved in selecting the optimal weights for each agent.

Algorithm 1 Multi-Agent SRL optimization method.

1: Initialize neural network of agent l as $\mu_l(s \mid w_{\mu_l})$
2: **for** $j = 1$ to M **do**
3:     Initialize random weights of each agent $w_{\mu_{l,M}}$ M times
4:     **for** $i = 1$ to l **do**
5:         Execute action a_{t_l}, observe reward r_{t_l} at each step t in common state s_t
6:         **if** Average reward $r_l > r_{\max_l}$ **then**
7:             Save weights of each agent $w_{\mu_{l,M}}$
8:         **end if**
9:     **end for**
10: **end for**

IV. NUMERICAL RESULTS

To test effectiveness of the proposed multi-agent strategy in coordinating an ANN-based controller, a control scheme was implemented in MATLAB/Simulink using the model parameters provided in Table I. The output power of the WTG and PV systems, as well as the load demand, were monitored and analyzed, as shown in Fig. 5.

The proposed controller was found to be effective in addressing the LFC problem, as shown in Fig. 6 (top plot), where it outperformed the FOPID controller. In terms of transient response, the ANN-based controller demonstrated faster response times with smaller overshoots and shorter settling times in a fractional order environment. The proposed approach also exhibited smaller steady-state errors compared to the traditional controller, as detailed in Table II. Our proposed algorithm yields lower values of integral absolute error (IAE), root mean square error (RMSE), and mean absolute error (MAE) when compared to the FOPID controller, as evidenced by the results. It is worth noting that for two FOPID controllers, the parameters are set to fixed values as $K_{p,DLC} = 100$, $K_{i,DLC} = 50$, $K_{d,DLC} = 0.5$, $\lambda_{DLC} = 1.35$, $\mu_{DLC} = 1.2$ and $K_{p,FC} = 1000$, $K_{i,FC} = 200$, $K_{d,FC} = 0.5$, $\lambda_{FC} = 1.2$, $\mu_{FC} = 1.25$. According to the results, we can see an insignificant influence of renewable energy sources disturbances on the proposed ANN controller. In contrast, the usual FOPID controllers show less flexibility in the distributive process. Here, we can see how ANN is robust and adjusts the

parameters of both controllers to optimal values after each time step.

To verify the effectiveness of the fractional order modeling strategy and its impact on the LFC enhancement compared to the classical integer order models, we repeated the experiment for the considered MG system for the case of integer order FC and DLC models with the same parameters. It can be seen in Fig. 6 (bottom plot) that the proposed method with using the same trained process for fractional order case, again shows better performance than the FOPID controller One may see that this resulted in a slightly degraded performance when compared to the fractional order case above. Nevertheless, in both cases the proposed controller was able to keep the frequency within the limits.

TABLE II
NUMERICAL RESULTS

System order	Controller	IAE	RMSE	MAE
Fractional	Proposed	**0.0221**	1.7678^{-5}	9.2214^{-6}
	FOPID	0.3125	1.7269^{-4}	1.3014^{-4}
Integer	Proposed	**0.5133**	4.1268^{-4}	2.1367^{-4}
	FOPID	0.6186	3.9259^{-4}	2.5763^{-4}
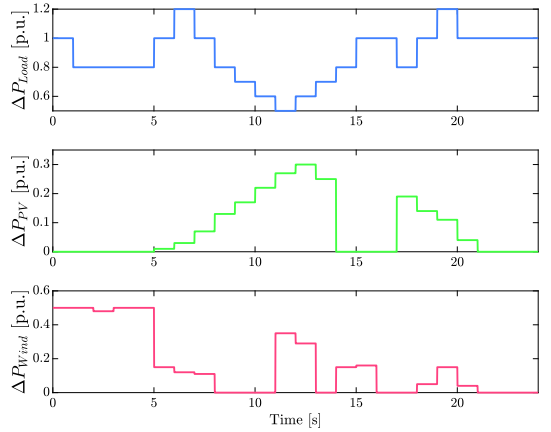
Fig. 5. Output power changes of renewable energy sources and load.

V. CONCLUSIONS

To overcome some issues related to the high-order nonlinear models in an isolated MG system, fractional order-based models were used to balance model accuracy and analytical complexity. In this study, a coordinated LFC strategy was proposed using two fractional order components to enhance the frequency stability. Two ANN controllers were deployed for each component in a multi-agent framework using a multi-agent SRL optimization technique, in addition were compared with two optimally tuned FOPID controllers. Test results

Fig. 6. Control performance of algorithms with fractional order (top) and integer order (bottom) models.

on an isolated MG with fractional components validated the effectiveness of the proposed coordinated LFC strategy. As illustrated in this study, it was observed t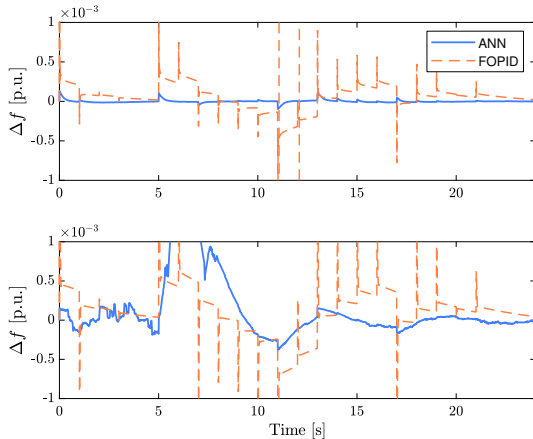hat while the proposed algorithm outperforms the FOPID controller in both systems with integer and fractional model of components, it exhibits an enhanced LFC system with greater stability in the MG system with fractional order subsystems.

REFERENCES

[1] J. M. Guerrero, J. C. Vasquez, J. Matas, L. G. De Vicuña, and M. Castilla, "Hierarchical control of droop-controlled AC and DC microgrids—a general approach toward standardization," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 1, pp. 158–172, 2010.

[2] M. Khan and M. Iqbal, "Dynamic modeling and simulation of a small wind–fuel cell hybrid energy system," *Renewable Energy*, vol. 30, no. 3, pp. 421–439, 2005.

[3] M. Nehrir, C. Wang, K. Strunz, H. Aki, R. Ramakumar, J. Bing, Z. Miao, and Z. Salameh, "A review of hybrid renewable/alternative energy systems for electric power generation: Configurations, control, and applications," *IEEE Transactions on Sustainable Energy*, vol. 2, no. 4, pp. 392–403, 2011.

[4] P. Thounthong, V. Chunkag, P. Sethakul, B. Davat, and M. Hinaje, "Comparative study of fuel-cell vehicle hybridization with battery or supercapacitor storage device," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 8, pp. 3892–3904, 2009.

[5] K. Jin, X. Ruan, M. Yang, and M. Xu, "Power management for fuel-cell power system cold start," *IEEE Transactions on Power Electronics*, vol. 24, no. 10, pp. 2391–2395, oct 2009.

[6] D.-J. Lee and L. Wang, "Small-signal stability analysis of an autonomous hybrid renewable energy power generation/energy storage system part I: Time-domain simulations," *IEEE Transactions on Energy Conversion*, vol. 23, no. 1, pp. 311–320, mar 2008.

[7] M. Uzunoglu and M. Alam, "Modeling and analysis of an FC/UC hybrid vehicular power system using a novel-wavelet-based load sharing algorithm," *IEEE Transactions on Energy Conversion*, vol. 23, no. 1, pp. 263–272, mar 2008.

[8] A. J. Veronica and N. S. Kumar, "Control strategies for frequency regulation in microgrids: A review," *Wind Engineering*, vol. 45, no. 1, pp. 107–122, aug 2019.

[9] A. Klem, M. H. Nehrir, and K. Dehghanpour, "Frequency stabilization of an islanded microgrid using droop control and demand response," in *North American Power Symposium (NAPS)*, 2016.

[10] A. K. Verma, H. B. Gooi, A. Ukil, N. R. Tummuru, and S. K. Kollimalla, "Microgrid frequency stabilization using model predictive controller," in *2016 IEEE PES Transmission Distribution Conference and Exposition-Latin America*, 2016.

[11] D. Oliveira, A. Z. de Souza, M. Santos, A. Almeida, B. Lopes, and O. Saavedra, "A fuzzy-based approach for microgrids islanded operation," *Electric Power Systems Research*, vol. 149, pp. 178–189, aug 2017.

[12] H. Bevrani, M. R. Feizi, and S. Ataee, "Robust frequency control in an islanded microgrid: H∞ and μ-synthesis approaches," *IEEE Transactions on Smart Grid*, pp. 1–1, 2015.

[13] A. Tepljakov, B. B. Alagoz, C. Yeroglu, E. A. Gonzalez, S. H. HosSeinia, E. Petlenkov, A. Ates, and M. Cech, "Towards industrialization of fopid controllers: A survey on milestones of fractional-order control and pathways for future developments," *IEEE Access*, vol. 9, pp. 21 016–21 042, 2021.

[14] M. Nayeripour, M. Hoseintabar, and T. Niknam, "Frequency deviation control by coordination control of FC and double-layer capacitor in an autonomous hybrid renewable energy power generation system," *Renewable Energy*, vol. 36, no. 6, pp. 1741–1746, jun 2011.

[15] K. Nosrati, H. Mansouri, and H. Saboori, "Fractional-order PID controller design of frequency deviation in a hybrid renewable energy generation and storage system," *CIRED - Open Access Proceedings Journal*, vol. 2017, no. 1, pp. 1148–1152, 2017.

[16] V. Skiparev, K. Nosrati, A. Tepljakov, E. Petlenkov, Y. Levron, J. Belikov, and J. M. Guerrero, "Virtual inertia control of isolated microgrids using NN-based VFOPID controller," *IEEE Transactions on Sustainable Energy*, 2023.

[17] K. Nosrati, A. Tepljakov, E. Petlenkov, Y. Levron, V. Skiparev, and J. Belikov, "Coordinated PI-based frequency deviation control of isolated hybrid microgrid: An online multi-agent tuning approach via reinforcement learning," in *2022 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*, 2022, pp. 1–5.

[18] H. Bevrani, F. Habibi, and S. Shokoohi, "ANN-based self-tuning frequency control design for an isolated microgrid," in *Meta-Heuristics Optimization Algorithms in Engineering, Business, Economics, and Finance*, P. Vasant, Ed. IGI Global, 2013, pp. 357–385.

[19] S. K. Akula and H. Salehfar, "Frequency control in microgrid communities using neural networks," *2019 North American Power Symposium (NAPS)*, pp. 1–6, 2019.

[20] Y. Wang and G. Zhao, "A comparative study of fractional-order models for lithium-ion batteries using runge kutta optimizer and electrochemical impedance spectroscopy," *Control Engineering Practice*, vol. 133, p. 105451, 2023.

[21] A. Tepljakov, E. Petlenkov, and J. Belikov, "FOMCON toolbox for modeling, design and implementation of fractional-order control systems," in *Applications in Control*, I. Petráš, Ed. De Gruyter, 2019, pp. 211–236.

[22] Z. Deng, H. Cao, X. Li, J. Jiang, J. Yang, and Y. Qin, "Generalized predictive control for fractional order dynamic model of solid oxide fuel cell output power," *Journal of Power Sources*, vol. 195, no. 24, pp. 8097–8103, 2010.

[23] T. J. Freeborn, B. Maundy, and A. S. Elwakil, "Fractional-order models of supercapacitors, batteries and fuel cells: a survey," *Materials for Renewable and Sustainable Energy*, vol. 4, pp. 1–7, 2015.

[24] B. Vedik, R. Kumar, R. Deshmukh, S. Verma, and C. K. Shiva, "Renewable energy-based load frequency stabilization of interconnected power systems using quasi-oppositional dragonfly algorithm," *Journal of Control, Automation and Electrical Systems*, vol. 32, no. 1, pp. 227–243, oct 2020.

[25] J. Sabatier, M. Merveillaut, J. M. Francisco, F. Guillemard, and D. Porcelatto, "Fractional models for lithium-ion batteries," in *2013 European Control Conference (ECC)*. IEEE, 2013, pp. 3458–3463.

[26] J. Tian, R. Xiong, W. Shen, and J. Wang, "A comparative study of fractional order models on state of charge estimation for lithium ion batteries," *Chinese Journal of Mechanical Engineering*, vol. 33, no. 1, pp. 1–15, 2020.

[27] Y. Wang, G. Gao, X. Li, and Z. Chen, "A fractional-order model-based state estimation approach for lithium-ion battery and ultra-capacitor hybrid power source system considering load trajectory," *Journal of Power Sources*, vol. 449, p. 227543, 2020.

Appendix 9

I
K. Nosrati, V. Skiparev, A. Tepljakov, E. Petlenkov, Y. Levron, and J. Belikov. 9 - application of neural network based variable fractional order pid controllers for load frequency control in isolated microgrids. In D. K. Mishra, L. Li, J. Zhang, and M. J. Hossain, editors, *Power System Frequency Control*, pages 203–216. Academic Press, 2023

Application of neural network based variable fractional order PID controllers for load frequency control in isolated microgrids

9

Komeil Nosrati[a], Vjatseslav Skiparev[b], Aleksei Tepljakov[a], Eduard Petlenkov[a], Yoash Levron[c], and Juri Belikov[b]
[a]Department of Computer Systems, Tallinn University of Technology, Tallinn, Estonia,
[b]Department of Software Science, Tallinn University of Technology, Tallinn, Estonia,
[c]The Andrew and Erna Viterbi Faculty of Electrical & Computer Engineering, Technion—Israel Institute of Technology, Haifa, Israel

9.1 Introduction

The problem of load frequency control (LFC) within prescribed nominal values is critical for the reliable operation of modern power systems [1]. The function of microgrids (MGs) should be studied more carefully, particularly when they are isolated and operate in remote locations. Because MGs in this case include converter-based renewable energy systems (RESs) and integrated energy storage systems (ESSs) with low inertia, which present additional challenges, more sophisticated control approaches are necessary to ensure uninterrupted load delivery with excellent power quality [2]. Even though several research studies have addressed MG management and control from a variety of viewpoints in recent decades, the LFC problem in an isolated MG is exacerbated by the rising penetration of RESs and ESSs, which necessitates additional system reliability considerations [3].

Wind turbine generator (WTG) and photovoltaic (PV) systems are two of the most promising RESs in isolated hybrid MGs (HMGs) [4]. Extra power generation components, such as fuel cells (FCs), are placed in the system to reduce power fluctuations caused by changing weather conditions and to provide a constant power supply [5]. Due to some drawbacks of the FCs, including slow dynamics and time delay resulted from its inherent physical properties, it must always be used in conjunction with ESSs to regulate power to a desired level and improve system performance [6]. The typical DC link capacitor cannot adapt to variations in load demand because of the constrained control operation of FC and its long-time delay. As a result, double-layered capacitors (DLCs) with fast power response can supplement the slower power output of the main source to compensate for the load variations and transient responses [7].

In an isolated HMG, the power for the load side demand can be efficiently generated and supplied by several energy generation sources with effective coordination of

frequency deviation control (FDC) among various subsystems [8]. Hitherto, multiple control methodologies for LFC in isolated HMG have been presented where the systems under study experienced various disturbances. Among different efforts, the small signal stability analysis of an HMG with an isolated load was explored in [9], in which the suggested system could achieve a proper frequency balance at different operating points. The load FDC of an isolated small-hydro plant was presented in [10], utilizing the reduced dump or resistive load approach and on/off control technique. A coordinated proportional-integral-differential (PID)-type control between FC and DLC systems was suggested in [11,12] for an isolated HMG as a contribution to FDC, with certain enhancements in quantitative and qualitative characteristics compared to prior studies.

In engineering practices, an efficient isolated HMG operation can be achieved by optimal parameter tuning of PID-type controllers. In this regard, several control and optimization procedures for the FDC of isolated HMG have been developed [13,14]. Very recently, tuned fractional order (FO)PID controllers were applied for automatic LFC problem of islanded HMG in a hybrid classical and advanced strategy [15,16]. However, all these offline parameter tuning efforts experienced several issues such as premature convergence, time complexity, and parameter tuning search space to obtain a higher degree of FDC. Applying the constant FOPID in these works might limit a control engineer from manipulating all characteristics of this controller, such as the usage of integration and derivative actions, toward a well-performing frequency control [17].

To overcome the LFC problems in the isolated HMG, we aim to provide the synthesis of a new flexible (FO)PID-based FDC to evaluate the grid stability under different levels of RESs penetration and load disturbances. For this goal, we describe an online tuning of the variable (FO)PID (V(FO)PID) controller using a neural network (NN) trained by the stochastic reinforcement learning (SRL) approach. This automatic approach, which admits capturing all the controller advantages with its additional features and "tuning knobs", such as improved robustness and disturbance rejection, as well as its contributions to time delay systems, can lead to the generation of such control laws that improve the performance of the control system. The research motivation in this chapter is to design an online multiagent method to improve the system performance by addressing the problem of FDC for an islanded HMG based on a self-tuned NN-V(FO)PID controller connected with SRL. It is worth noting that the presented approach demonstrates a successful controller tuning effort in a multiagent structure, which is valid for FDC in isolated HMG typologies.

9.2 Isolated HMG configuration and mathematical modeling

9.2.1 System structure

A typical isolated power generation and energy storage system — the isolated HMG under investigation — is shown in Fig. 1. The isolated HMG comprises several power generation and storage subsystems referred to as WTG, PV, FC, and the DLC bank. To provide isolated loads, all these four subsystems are linked in parallel to a shared AC
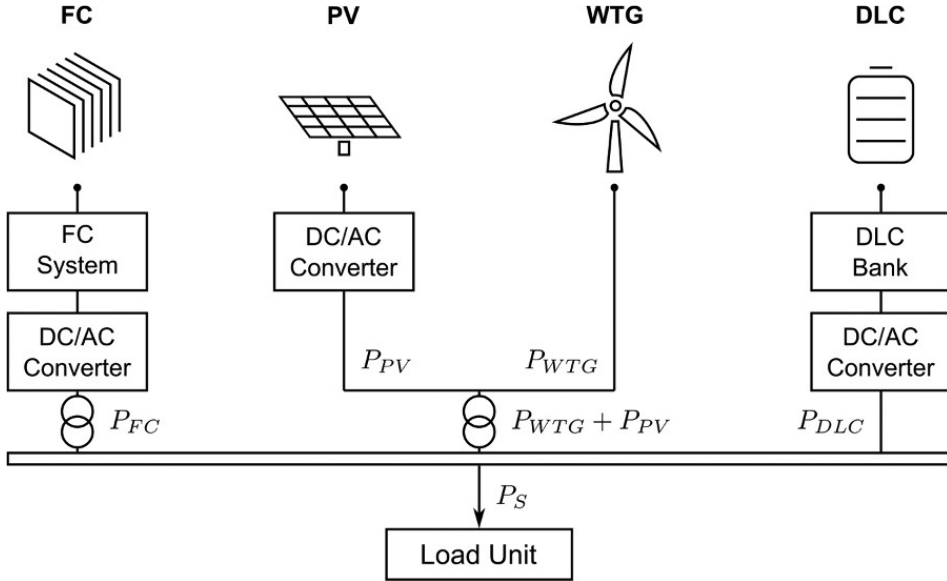
Fig. 1 A schematic diagram of the isolated HMG.

bus line, with PV, FC, and DLC connected via three independent DC/AC converters. This system uses the PV and WTG as major energy sources, with power generation taking precedence to fulfill the load demand. The dynamic characteristic of the system and components are usually time-varying with high order nonlinear dynamics. However, linearized low-order models are employed to execute LFC synthesis concerning the variations in load or output power of RESs.

9.2.2 WTG model

Mechanical power in the turbine section is dictated by wind speed, which has a direct impact on WTG output power P_{WTG}. This effect can be illustrated as $P_{WTG} = 0.5\rho A v^3 C_p(\lambda, \theta)$, where ρ (kg/m^3) is the density of air, A (m^2) is the area swept by the blades of the rotor, v (m/s) is the wind velocity, and C_p is the rotor efficiency, which is a function of tip speed ratio λ and pitch angle θ. The maximum value of C_p can be obtained for a particular direction of the blades and when the λ is at its specific value (see Fig. 2A). The rotor speed can preserve λ at its optimal value depending on v, allowing the maximum energy from the wind to be utilized. Fig. 2B explains the variation in P_{WTG} as a function of v. By control of θ, the power stayed constant when v increased through the rated wind velocity and takes its maximum value when $14 < v < 25$. For $v > 25$, the system shuts down. By introducing T_{WTG} as the time constant of the turbine, the mathematical model of WTG can be given by

$$G_{WTG}(s) = \frac{\Delta P_{WTG}}{\Delta P_{wind}} = \frac{1}{sT_{WTG} + 1}, \tag{1}$$

where ΔP_{wind} and ΔP_{WTG} are the variations of the WTG mechanical and output power, respectively.
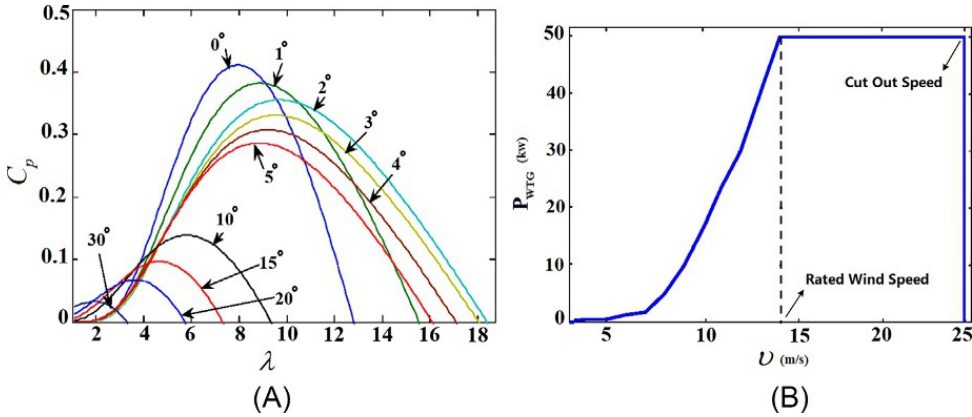
Fig. 2 Variation of C_p and P_{WTG} as a function of tip speed ratio and wind velocity [11]. (A) C_p–λ and (B) P_{WTG}–υ.

9.2.3 PV model

The PV system, which comprises solar panels coupled in series and parallel structure, is one of the most promising, adaptable, and environmentally friendly power sources. Based on solar cell temperature and array area, PV transforms solar radiation to electrical data. The output power P_{PV} can be defined in terms of conversion efficiency η (°C^{-1}), PV array area S (m^2), solar irradiation φ (W/m^2), and ambient temperature T_a (°C) as $P_{PV} = \eta S \varphi (1 - 0.005(T_a - 25))$. Variations in solar irradiation to PV system output power can be characterized by

$$G_{PV}(s) = \frac{\Delta P_{PV}}{\Delta \varphi} = \frac{1}{sT_{PV} + 1}, \tag{2}$$

where T_{PV} is the PV time constant. Due to dust accumulation and temperature variations, the real efficiency of module might be as low as 70% of the standard test conditions efficiency given by the manufacturer.

9.2.4 FC model

RESs that incorporate the DLC system are appropriate for power supply stabilization. The suggested system, on the other hand, is mostly made up of natural energy sources including FC technology, which is regarded as a high-efficiency power generation system. The fuel supply portions of this static device, such as pumps and valves, have slow dynamics, resulting in decreased power production. The first-order time delay transfer function

$$G_{FC}(s) = \frac{1}{sT_{FC} + 1}, \tag{3}$$

with T_{FC} as the FC time constant, is used to calculate the dynamic response of the FC system. To simulate and explore our isolated HMG with complex components, fractional-order models can also be used in FC and DLC modelling which offers the benefit of improved accuracy as compared to traditional integer models, especially for the dynamics under transient operating circumstances [18–20].

9.2.5 DLC model

Electrochemical capacitors with porous electrodes and the DLC charging process play a critical role in meeting the demands of a variety of applications, including electric cars, power quality, and ESS technology for RESs. DLCs, sometimes known as super-capacitors or ultracapacitors, can be represented by

$$G_{DLC}(s) = \frac{1}{sT_{DLC} + 1}, \tag{4}$$

where T_{DLC} is the time constant, with bigger frequency bands needing a higher number of parameters.

9.2.6 Frequency deviation model

An effective control of given subsystems is required for the isolated HMG in its steady operation with various power generators. This is accomplished by controlling the variation in the frequency profile Δf, which can be given by $\Delta f = K_{HMG}^{-1}\Delta P$, with $\Delta P = P_{net} - P_{load}$, where P_{net} and P_{load} are the net and load power, respectively, and K_{HMG} is the system frequency constant of the given HMG. The dynamical model of frequency variation based on per unit power deviation may be expressed as follows:

$$\frac{\Delta f}{\Delta P} = \frac{1}{sM + D}, \tag{5}$$

where M and D are inertia and damping constants of the isolated HMG, respectively.

9.3 (FO)PID controllers, actions, and tuning rules

PID-type controllers have a wide range of popularity due to their ease of design and strong performance, which includes low overshoot and short settling time for slow processes. (FO)PID controllers are less sensitive to parameter changes and can easily achieve the property of iso-damping. The general form of a constant (FO)PID controller is as

$$y(t) = K_p e(t) + K_i D^{-\lambda} e(t) + K_d D^{\mu} e(t), \tag{6}$$
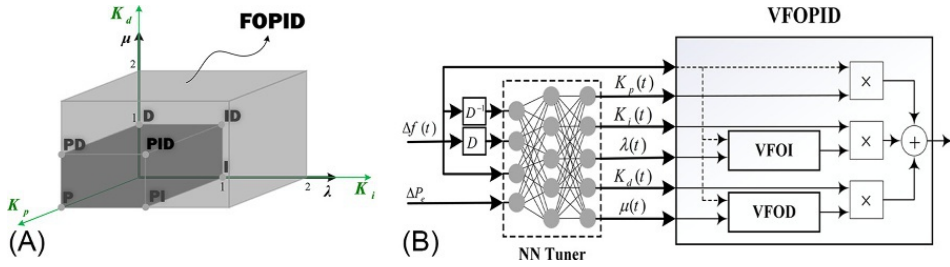
Fig. 3 The proposed V(FO)PID controller: (A) basic scheme and (B) its NN-based tuning scheme.

where D^α is fractional derivative and integral operators for positive and negative α, respectively, and the positive real constants μ and λ are the integration and differentiation orders, respectively. Also, $u(t)$ and $e(t) = \Delta f(t)$ are control and error signals, respectively, and K_p, K_i, and K_d are the constant gains. Although this controller, when the orders are not equal to one, provides additional tuning freedom, allowing it to meet precise control performance, it can bound a control engineer to manipulate all the features. Instead, the V(FO)PID controller with five online tunable parameters can considerably improve the control performance requirements. According to Fig. 3A, we can not only move continuously in the PID plane instead of jumping between the fixed points but we can also search for desirable controller parameters in a space inside the cube and between the eight vertices.

9.3.1   Control actions

When the related gains and orders of each function change, both integration and differential functions have an impact on the steady-state process and dynamic features. In general, a PID-type controller has three control actions:

- *Proportional action*: When the control error is minor, the proportional advantage is its ability to supply a smaller control variable and minimize unnecessary control efforts.
- *Integral action*: While a large K_i will make a system more unstable with oscillations, a small gain will cause the system to deviate from its optimal dynamic performance. The frequency band is also large for small λ, and the system is stable with quick reaction and static error. In the other direction, an oversize of λ will overshadow the system stability with the increase of the overshoot, rise time, and settling time.
- *Derivative action*: K_d has no effect on steady-state error, but may enhance dynamic features. With a small K_d, the overshoot and settling time will rise, but with a large gain, the system noise will grow, and the system performance in disturbance tolerance will be degraded. When μ is small, it increases the response accuracy. However, increasing this order reduces the system overshoot and settling time, and for an oversize μ, the closed-loop stability will be negatively affected.

To achieve sufficient control performance for a target system, it is required to keep not only the gains but also the fractional orders of the controllers in a desirable range. This can be obtained by optimal tuning of the parameters of the PID-type controllers toward efficient operation of the system.

9.3.2   Tuning rules

The tuning of most PID-type controllers is always a challenging task. According to the literature, there are three types of tuning procedures for PID-type controllers [21]: heuristic, rule-based, and model-based tuning approaches. A heuristic tuning approach, such as the trial-and-error method, is one that uses general rules to provide approximate or qualitative outcomes. This is currently the most often used method in industrial control; however, due to the obvious drawbacks it is not acceptable to use this method for intelligent system control where the goal is to achieve the best possible control system performance while minimizing the required control effort to help reduce energy waste. On the other hand, by using simple mathematical methods, rule-based approaches, such as Ziegler-Nichols and Cohen-Coon, can tune a PID-type controller by assuming a specific process response. Model-based tuning, also known as optimization-based tuning, helps one to achieve the parameters optimally. In addition to these three methods, offline and online methods are two other well-known categories in the tuning of PID-type controllers. Unlike offline tuning, the online method admits capturing all the controllers' advantages with its additional features, resulting in improved performance and disturbance rejection. For online tuning purposes, the V(FO)PID controller can be given as

$$y(t) = K_p(t)e(t) + K_i(t)D^{-\lambda(t)}e(t) + K_d(t)D^{\mu(t)}e(t), \tag{7}$$

where $D^{-\lambda(t)}$ and $D^{\mu(t)}$ are variable fractional order integral (VFOI) and derivative (VFOD), respectively. In the case of using fractional operators, a powerful computing device is required to implement the controller (7) with fixed- or floating-point arithmetic. Therefore, we apply parallel connections of fractional order modeling and control (FOMCON) library blocks proposed in [22] to offer a real-time switch between fractional operators with an order resolution of 0.1. These functions switch between multiple fractional integration and differentiation operators and are used to implement a VFOPID controller according to the tuned parameters in combination with switching objective functions and emerging from an input signal produced from the NN tuner system (see Fig. 3B).

9.4   The proposed (FO)PID-based LFC: Multiagent NN-based online tuning approach

9.4.1   Coordinated control strategy

The FDC problem in the proposed isolated HMG can be handled with improved power quality by utilizing a coordinated control approach between FC and DLC subsystems, see Fig. 4. These two devices serve as backups, compensating for large and low frequency deviations, respectively. Also, in long-term operation, a high-pass filter (HPF) can decrease DLC charging and discharging. The net power generation is determined by

$$P_{net} = P_{WTG} + P_{PV} + P_{FC} \pm P_{DLC}. \tag{8}$$
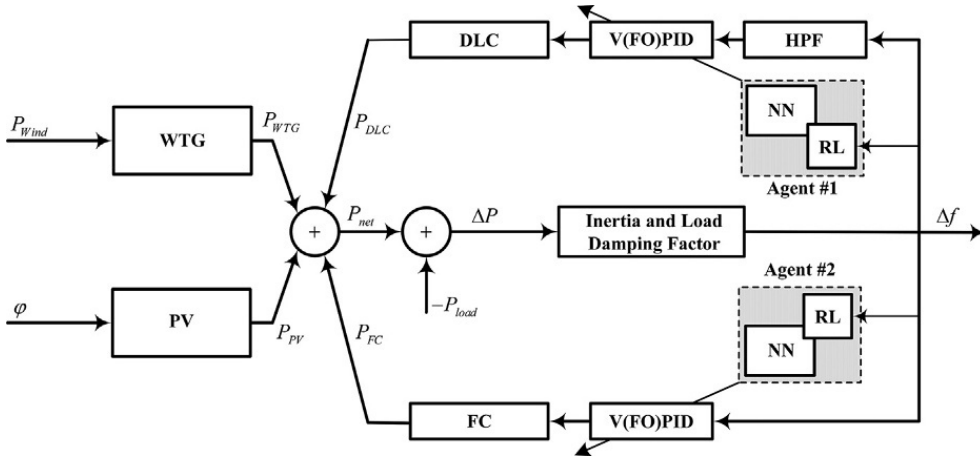
Fig. 4 Proposed coordinated multiagent FDC scheme.

For the control purpose, an online tuning-based V(FO)PID controller is used, which admits capturing all notable features of the controller. It is worth mentioning that the resilience of the isolated HMG can be increased by online tuning of the controller using NN technique learned by SRL algorithm, which select right combination of NNs parameters to avoid conflict between agents and realize which V(FO)PID coefficients are more significant. This approach of selecting the parameters is beneficial in FDC purposes in the HMG, which necessitates additional system reliability considerations. In this regard, a multiagent structure is used to adjust the controller, which is beneficial for the FDC purpose. As illustrated in the next sections, our online-tuned NN-based V(FO)PID controller technique makes a successful effort toward FDC in the given isolated HMG.

9.4.2 NN-based online tuner

The main idea of NN-based online tuner is an architecture design that uses the control error signal to change the values of the V(FO)PID parameters [23]. The PID-type controller can be automatically adapted to any process using this intelligent technique. On the other hand, it may be difficult to choose an appropriate architecture of the NN-based tuner. In the proposed approach, evolutionary algorithms, such as the neuro-evolutionary of augmenting topologies (NEAT) algorithm, tackle this problem by changing the topology of NNs. We propose that the NN-based tuner adhere to the PID-type controller concept, i.e., the NN-based tuner must include an integrator D^{-1} and an optional differential element D due to perturbations in the control loop which can destabilize the tuning process of all controller coefficients, see Fig. 4. Moreover, NN must be built in such a way that negative outputs are avoided. Nevertheless, since variable $K_d(t)$ is very sensitive to any rapid change of controller error, the V(FO)PI version is the better choice for most processes. In the case of variable orders, we must be sure about exact the range and resolution of outputs because they are computationally complex in real-time control.

Selection of the suitable activation function is very important in the design of an NN-based tuner, where it must transform all negative inputs to a positive output. Two original *tansig* and *sigmoid* activation functions are not acceptable due to the negative range of the first function and a failure to address the rising output for negative inputs. As a solution, we propose the following modified *tansig* function

$$y = \frac{n}{1 + e^{-|x|}} - \frac{n}{2} \tag{9}$$

with absolute input value of x and constrained output range determined by constant n. Here, the absolute value of x is used to avoid negative output, provide a sufficient range of coefficients for controller, and make the tuning more robust. We avoid connecting the HPF to the NN-based tuner since it has a negative impact on the stability of the tuner and confuses the reward system. Therefore, the input signal for both agents is the same.

9.4.3   SRL-based training for multiple agents

Using RL technique and in a combined strategy of both heuristic and rule-based methods, we try to obtain the optimal NN weights. In our proposed method, the effectiveness of the NN training system associated with RL is largely determined by the design of the rewarding mechanism. In FDC purposes of the isolated HMG from any order and degree of complexity, the most significant feature is the allowable range of frequency deviation. However, the technique for designing a standard reward system remains still an open question, as each application necessitates a unique approach. Furthermore, the optimal range of weights for every tuner is also individual, but here due to the unlimited range of optimal controller coefficients, it is not clear how machine learning (ML) methods can find weights with right magnitude. To avoid some problems related to the optimization algorithms, such as gradient descent in which it takes a long time to reach a local minimum and only allows for a narrow range of ideal weights, here we propose SRL method for effective tuning of the parameters.

To train numerous RL agents and enable optimal online tuning of each V(FO)PID controller, we employ a basic SRL in which random weights are generated M times. The training is based on reward feedback arranged by the reward system provided in [22], which allows for the best weights by selecting actions with the highest potential reward. In this method, we perform actions at a_t (i.e., $K_p(t)$, $K_i(t)$, and $K_d(t)$) and check the received maximum average reward $r_{avr,\max}$ of each agent individually, where the main criterion for rewarding is the magnitude of Δf (see (10)), and an agent receives a positive reward $+r$ if Δf is less than 0.01 Hz; otherwise, it receives a negative reward.

$$r_t = \begin{cases} (0.5 + 10|\Delta f|)^{-1}, & \text{if } 10|\Delta f| < 0.01, \\ -20|\Delta f|, & \text{if } 10|\Delta f| > 0.01. \end{cases} \tag{10}$$

9.5   Simulation results

Dynamic models for the main subsystems were constructed using MATLAB/ Simulink to develop an overall power management strategy for the proposed system and to examine its performance. The parameters used in the system modeling are as $T_{WTG}=1.5$ s, $T_{PV}=1.8$ s, $T_{FC}=0.26$ s, $T_{DLC}=0.01$ s, $M=0.4$, and $D=0.03$. Also, Fig. 5A and B illustrates the real wind speed and sun irradiation, respectively. The time scale utilized in analysis is expressed by sampling time, while the time on the representative day to be simulated is expressed by assumption time. In simulation, the sampling time interval was chosen as 5×10^{-5}, which equals to 0.003 min in the assumption time. The power produced by WTG and PV systems over the course of a day is shown in Fig. 5C. From this figure, step load demands are applied to the system to demonstrate effectiveness of the proposed method.
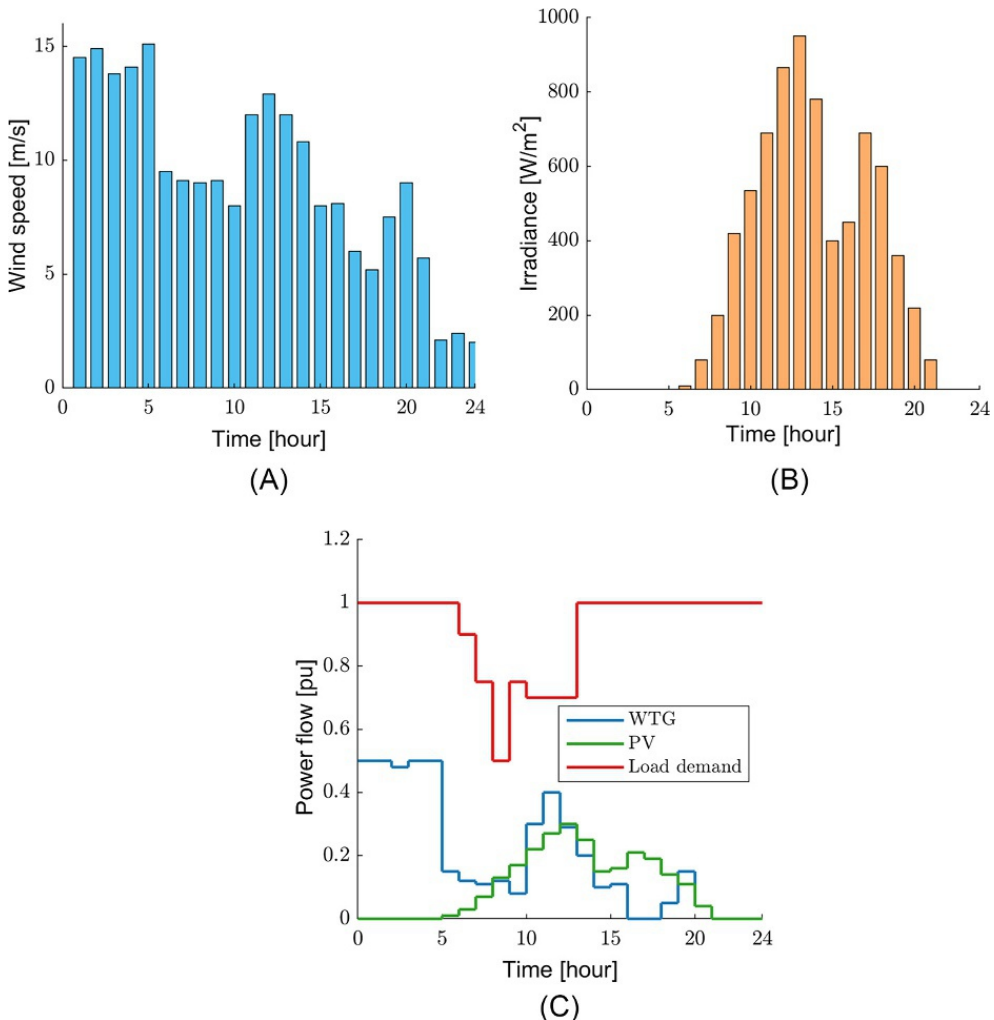

Fig. 5 (A, B) The real data from RESs and (C) power produced by WTG and PV and variation of load demand.
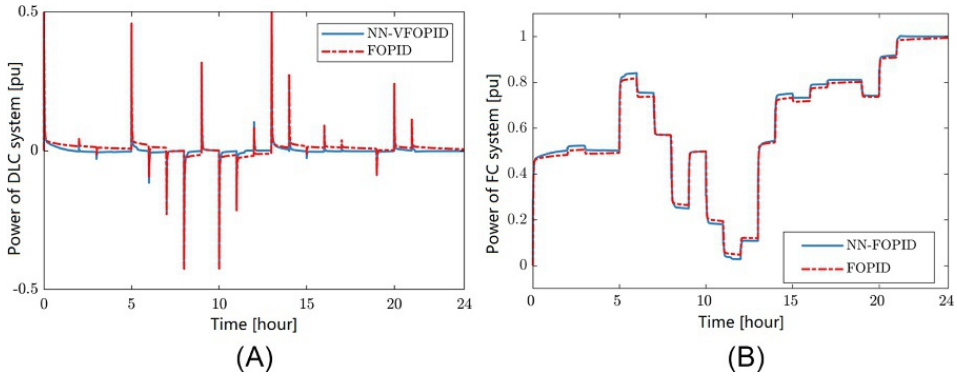
Fig. 6 (A) DLC output power and (B) FC output power with and without NN.

Fig. 6A and B shows the output power of DLC and FC, respectively, with and without NN. As seen, the lower required ESS capacity can be found by using the NN-FOPID. As shown in Fig. 7, the frequency deviation can be controlled suitably by coordinating FC and DLC to compensate for the shortage and complement full hybrid power generation with considering the impact of system frequency variation. According to the results, we can see an insignificant influence of WTG and PV disturbances on the proposed combination of NN and variable PID-type controller. In contrast, the traditional controllers show less flexibility in this case. In terms of transient response, the NN-based V(FO)PID controller has a smaller overshoot with a shorter settling time, thus indicating a faster transient time and hence better reference tracking performance. In terms of steady state, the NN-based controllers have small steady-state errors and reduce the values of the integral absolute error (IAE), mean square error (MSE), and root MSE (RMSE) when compared to the traditional PID-type controller (see Table 1). As seen, the proposed combination of NN and V(FO) PID produces the best results. One can see that when NN is combined with V(FO) PID, the reactions to the system disturbances are reduced. Fig. 8A and B illustrates parameters of the online tuned V(FO)PID controller for both agents 1 and 2. Here, we can see how NN adjusts the parameters of both controllers to optimal values after each step in the signal.
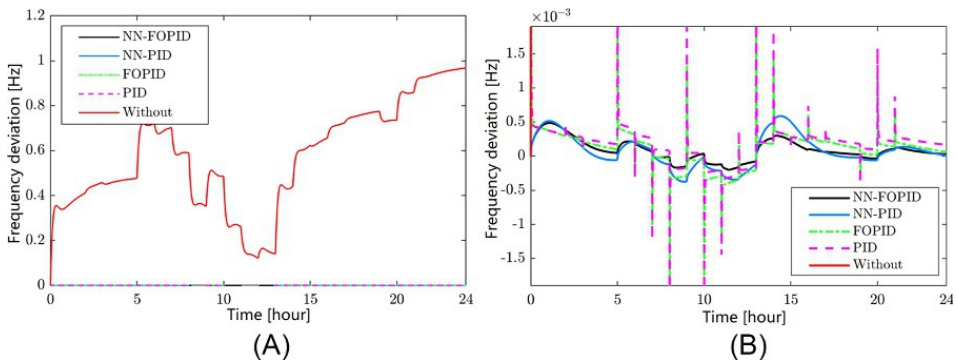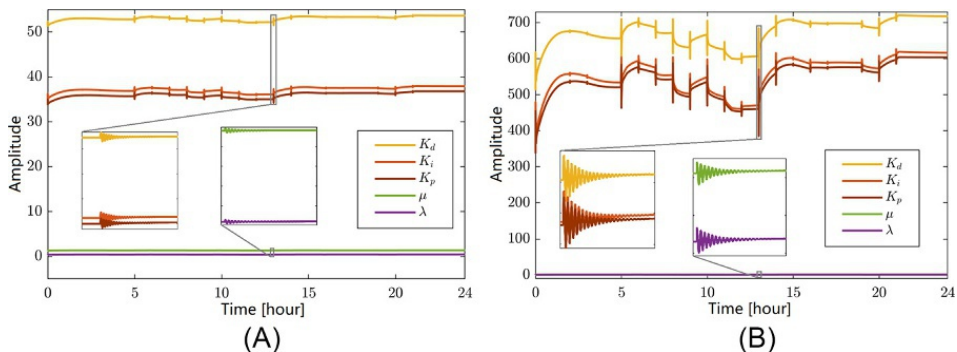

Fig. 7 Frequency deviation: (A) standard and (B) zoom in modes.

Table 1 Performance result of isolated HMG with different controllers.

	NN-FOPID	NN-PID	FOPID	PID	Without
IAE	0.003324	0.004559	0.005774	0.006403	13.78
RMSE	0.0001804	0.0002482	0.000322	0.000323	0.6201
MSE	3.2528×10^{-8}	6.1618×10^{-8}	1.0633×10^{-7}	1.0416×10^{-7}	0.3845

Fig. 8 Online tuning of VFOPID controllers: (A) agent 1 and (B) agent 2.

9.6 Conclusion

In this chapter, we presented a novel combination of NN and V(FO)PID controllers applied to the LFC problem in isolated HMG with high RES penetration, where the SRL algorithm optimized all tuned parameters (both gains and orders) of this controller. We demonstrated that combining a V(FO)PID controller with an NN-based tuner is an effective strategy for FDC tasks. Unlike traditional PID-type controllers, the NN-based V(FO)PID controller has active support from a ML algorithm, allowing it to self-adapt to HMG disturbances and produce smooth frequency deviation. The results demonstrate that the proposed combination of PID-type controllers and NN tuner captures all the remarkable advantages of the controller and gives rise to the generation of control laws that improve both transient and steady-state errors. In summary, the major benefits of the proposed solution can be summarized as follows:

- An effective NN-tuning-based system for V(FO)PID controller coefficients using SRL strategy to find optimal weights of the controller.
- Fast training of NN-based tuner by the proposed effective SRL algorithm.
- Capturing all tuning knobs of the V(FO)PID controller by designing a self-tuning technique, resulting in robustness improvement, fast recovery time, and frequency stability enhancement.
- Capability of the designed controller over a wide range of operating conditions due to its flexibility in use of integration and derivative actions toward a well performance in frequency support operations with multiple DGs and RESs.

High precision modeling of the subsystems utilizing mathematical methodologies such as fractional calculus, notably generating and storage elements like FCs and DLCs, will be the foundation for the future study and development of the HMG, particularly for the estimation and energy management.

Acknowledgments

References

[1] S. Kayalvizhi, D.V. Kumar, Load frequency control of an isolated micro grid using fuzzy adaptive model predictive control, IEEE Access 5 (2017) 16241–16251.

[2] D.E. Olivares, A. Mehrizi-Sani, A.H. Etemadi, C.A. Cañizares, R. Iravani, M. Kazerani, A.H. Hajimiragha, O. Gomis-Bellmunt, M. Saeedifard, R. Palma-Behnke, G.A. Jiménez-Estévez, Trends in microgrid control, IEEE Trans. Smart Grid. 5 (4) (2014) 1905–1919.

[3] R. Majumder, A. Ghosh, G. Ledwich, Load frequency control in a microgrid: Challenges and improvements, in: Smart Power Grids, Springer, Berlin, Heidelberg, 2012, pp. 49–82.

[4] O. Kraa, H. Ghodbane, R. Saadi, M.Y. Ayad, M. Becherif, A. Aboubou, M. Bahri, Energy management of fuel cell/supercapacitor hybrid source based on linear and sliding mode control, Energy Procedia 74 (2015) 1258–1264.

[5] M. Uzunoglu, O.C. Onar, M.S. Alam, Modeling, control and simulation of a PV/FC/UC based hybrid power generation system for stand-alone applications, Renew. Energy 34 (3) (2009) 509–520.

[6] P. Thounthong, V. Chunkag, P. Sethakul, S. Sikkabut, S. Pierfederici, B. Davat, Energy management of fuel cell/solar cell/supercapacitor hybrid power source, J. Power Sources 196 (1) (2011) 313–324.

[7] P.G. Arul, V.K. Ramachandaramurthy, R.K. Rajkumar, Control strategies for a hybrid renewable energy system: a review, Renew. Sustain. Energy Rev. 42 (2015) 597–608.

[8] P. Thounthong, A. Luksanasakul, P. Koseeyaporn, B. Davat, Intelligent model-based control of a standalone photovoltaic/fuel cell power plant with supercapacitor energy storage, IEEE Trans. Sustain. Energy 4 (1) (2012) 240–249.

[9] D.J. Lee, L. Wang, Small-signal stability analysis of an autonomous hybrid renewable energy power generation/energy storage system part I: Time-domain simulations, IEEE Trans. Energy Convers. 23 (1) (2008) 311–320.

[10] S. Doolla, T.S. Bhatti, Load frequency control of an isolated small-hydro power plant with reduced dump load, IEEE Trans. Power Syst. 21 (4) (2006) 1912–1919.

[11] M. Nayeripour, M. Hoseintabar, T. Niknam, Frequency deviation control by coordination control of FC and double-layer capacitor in an autonomous hybrid renewable energy power generation system, Renew. Energy 36 (6) (2011) 1741–1746.

[12] K. Nosrati, H.R. Mansouri, H. Saboori, Fractional-order PID controller design of frequency deviation in a hybrid renewable energy generation and storage system, CIRED-Open Access Proc. J. 1 (2017) 1148–1152.

[13] A. Rahman, L.C. Saikia, N. Sinha, Automatic generation control of an interconnected two-area hybrid thermal system considering dish-stirling solar thermal and wind turbine system, Renew. Energy 105 (2017) 41–54.

[14] J. Mudi, C.K. Shiva, B. Vedik, V. Mukherjee, Frequency stabilization of solar thermal-photovoltaic hybrid renewable power generation using energy storage devices, Iranian J. Sci. Technol., Trans. Electr. Eng. 45 (2) (2021) 597–617.

[15] A.X. Irudayaraj, N.I. Wahab, M.G. Umamaheswari, M.A. Radzi, N.B. Sulaiman, V. Veerasamy, S.C. Prasanna, R. Ramachandran, A Matignon's theorem based stability analysis of hybrid power system for automatic load frequency control using atom search optimized FOPID controller, IEEE Access. 8 (2020) 168751–168772.

[16] M.R. Chen, G.Q. Zeng, Y.X. Dai, K.D. Lu, D.Q. Bi, Fractional-order model predictive frequency control of an islanded microgrid, Energies 12 (1) (2018) 84.

[17] A. Tepljakov, B.B. Alagoz, C. Yeroglu, E.A. Gonzalez, S.H. Hosseinnia, E. Petlenkov, A. Ates, M. Cech, Towards industrialization of FOPID controllers: A survey on milestones of fractional-order control and pathways for future developments, IEEE Access. 9 (2021) 21016–21042.

[18] A. Choudhury, H. Chandra, A. Arora, Application of solid oxide fuel cell technology for power generation—a review, Renew. Sustain. Energy Rev. 20 (2013) 430–442.

[19] Y. Wang, M. Li, Z. Chen, Experimental study of fractional-order models for lithium-ion battery and ultra-capacitor: modeling, system identification, and validation, Appl. Energy 278 (2020), 115736.

[20] M. Yu, Y. Li, I. Podlubny, F. Gong, Y. Sun, Q. Zhang, Y. Shang, B. Duan, C. Zhang, Fractional-order modeling of lithium-ion batteries using additive noise assisted modeling and correlative information criterion, J. Adv. Res. 25 (2020) 49–56.

[21] D. Valerio, J.S. da Costa, A review of tuning methods for fractional PIDs, in: 4th IFAC Workshop on Fractional Differentiation and its Applications, 10(5), 2010.

[22] A. Tepljakov, E. Petlenkov, J. Belikov, I. Petráš, FOMCON toolbox for modeling design and implementation of fractional-order control systems, in: Applications in Control Berlin, De Gruyter, Germany, 2019, pp. 211–236.

[23] V. Skiparev, J. Belikov, E. Petlenkov, Reinforcement learning based approach for virtual inertia control in microgrids with renewable energy sources, in: 2020 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe), 2020, pp. 1020–1024.

Curriculum Vitae

1. Personal data

Name	Vjatšeslav Škiparev
Date and place of birth	15.11.1992 Sillamäe, Estonia
Nationality	Estonian

2. Contact information

Address	Tallinn University of Technology, School of Information Technology, Department of Software Science, Akadeemia Tee 21/1, 12618 Tallinn, Estonia
Phone	+372 58161833
E-mail	vjatseslav.skiparev@ttu.ee

3. Education

2019–2023	Tallinn University of Technology, School of Information Technology, Information and Communication Technology, PhD studies
2016–2018	Tallinn University of Technology, Faculty of Information Technology, Computer and Systems Engineering , MSc
2012–2016	Tallinn University of Technology, Virumaa College, Industrial Automation, BSc

4. Language competence

Estonian	fluent
English	fluent
Russian	native

5. Professional employment

2019–2023	Tallinn University of Technology, Early Stage Researcher
2018–2019	Virumaa Colleague of Tallinn University of Technology, Lecturer

6. Voluntary work

2013–2014	TUT Virumaa College Student Council

7. Computer skills

- Operating systems: Windows, Linux Ubuntu

- Document preparation: TeXStudio, Microsoft Office Word, JabRef, InkScape

- Programming languages: C/C++, MATLAB, Ladder, Python

- Scientific packages: MATLAB/Simulink, PSIM

8. Defended theses

- 2018, Control of Water Tank by NN-SANARX based Controller Implemented on Arduino–MATLAB/Simulink Platform, MSc, supervisor Prof. Eduard Petlenkov, Tallinn University of Technology, Faculty of Information Technology

- 2016, Modernization of semi-coke gas burner control system, BSc, supervisor Dr. Olga Ruban, Tallinn University of Technology, Virumaa College

9. Field of research

- Power systems

- Load-Frequency Control

- Process Control

- Artificial neural networks

- Reinforcement learning

10. Scientific work
Papers

1. V. Skiparev, J. Belikov, and E. Petlenkov. Reinforcement learning based approach for virtual inertia control in microgrids with renewable energy sources. In *IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe)*, The Hague, NL, 2020

2. V. Skiparev, J. Belikov, and E. Petlenkov. MIMO reinforcement learning based approach for frequency support in microgrids with high renewable energy penetration. In *IEEE PES General Meeting*, Washington DC, USA, 2021

3. V. Skiparev, R. Machlev, N. R. Chowdhury, Y. Levron, E. Petlenkov, , and J. Belikov. Virtual inertia control methods in islanded microgrids. *Energies*, 14(6):1562, 2021

4. K. Nosrati, V. Skiparev, A. Tepljakov, E. Petlenkov, Y. Levron, and J. Belikov. Coordinated PI-based frequency deviation control of isolated hybrid microgrid: An online multi-agent tuning approach via reinforcement learning. In *2022 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*, pages 1–5, 2022

5. V. Skiparev, J. Belikov, E. Petlenkov, and Y. Levron. Reinforcement learning based MIMO controller for virtual inertia control in isolated microgrids. In *IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*, Novi Sad, Serbia, 2022

6. V. Skiparev, K. Nosrati, A. Tepljakov, E. Petlenkov, Y. Levron, J. Belikov, and J. M. Guerrero. Virtual inertia control of an isolated microgrid using NN-VFOPID controller: A self-tuning approach. *IEEE Transactions on Sustainable Energy*, 2023

7. K. Nosrati, V. Skiparev, A. Tepljakov, E. Petlenkov, J. Belikov, and Y. Levron. Constrained intelligent frequency control in an AC microgrid: An online reinforcement learning based PID tuning approach. In *2023 IEEE Power and Energy Society General Meeting*, pages 1–5, 2023

8. V. Skiparev, K. Nosrati, J. Belikov, A. Tepljakov, and E. Petlenkov. An enhanced NN-based load frequency control design of MGs: A fractional order modeling method. In *2023 IEEE International Conference on Compatibility, Power Electronics and Power Engineering (CPE-POWERENG)*, Tallinn, Estonia, 2023

Elulookirjeldus

1. Isikuandmed

Nimi	Vjatšeslav Škiparev
Sünniaeg ja -koht	15.11.1992, Sillamäe, Eesti
Kodakondsus	Eesti

2. Kontaktandmed

Aadress	Tallinna Tehnikaülikool, Tarkvarateadus Instituut, Ehitajate tee 5, 19086 Tallinn, Estonia
Telefon	+372 58161833
E-post	vjatseslav.skiparev@taltech.ee

3. Haridus

2019–2023	Tallinna Tehnikaülikool, infotehnoloogia teaduskond, Info- ja kommunikatsioonitehnoloogia, doktoriõpe
2016–2018	Tallinna Tehnikaülikool, infotehnoloogia teaduskond, Arvutisüsteemid, MSc
2012–2016	Tallinna Tehnikaülikool, Virumaa Kolledž, Töötmise Automatiseerimine, BSc

4. Keelteoskus

eesti keel	kõrgtase
inglise keel	kõrgtase
vene keel	emakeel

5. Teenistuskäik

2019–2023	Tallinna Tehnikaülikool, doktorant-nooremteadur
2018–2019	Tallinna Tehnikaülikool Virumaa Kolledž, lektor

6. Vabatahtlik töö

2013–2014	TTÜ Virumaa Kolledž Üliõpilasnõukogu

7. Computer skills

- Operatsioonisüsteemid: Windows, Linux Ubuntu

- Kontoritarkvara: TeXStudio, JabRef, Microsoft Office Word, InkScape

- Programmeerimiskeeled: C/C++, MATLAB, Ladder, Python

- Teadustarkvara paketid: MATLAB/Simulink, PSIM

8. Kaitstud lõputööd

- 2018, NN-ANARX mudelil põhinev veepagi juhtimissüsteem Arduino-Matlab/Simulink platvorm, MSc, juhendaja Prof. Eduard Pelenkov ja Prof. Juri Belikov, Tallinna Tehnikaülikool, Arvutisüsteemide Instituut

- 2016, Poolkooksigaasi põleti juhtimissüsteemi moderniseerimine, BSc, juhendaja Dr. Olga Ruban, Tallinna Tehnikaülikool, Virumaa Kolledž

9. Teadustöö põhisuunad

- Tehisintelekt
- Stiimulõpe
- Jühtimisüsteemid
- Energisüstemid

10. Teadustegevus
Publikatsioonid

1. V. Skiparev, J. Belikov, and E. Petlenkov. Reinforcement learning based approach for virtual inertia control in microgrids with renewable energy sources. In *IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe)*, The Hague, NL, 2020

2. V. Skiparev, J. Belikov, and E. Petlenkov. MIMO reinforcement learning based approach for frequency support in microgrids with high renewable energy penetration. In *IEEE PES General Meeting*, Washington DC, USA, 2021

3. V. Skiparev, R. Machlev, N. R. Chowdhury, Y. Levron, E. Petlenkov, , and J. Belikov. Virtual inertia control methods in islanded microgrids. *Energies*, 14(6):1562, 2021

4. K. Nosrati, V. Skiparev, A. Tepljakov, E. Petlenkov, Y. Levron, and J. Belikov. Coordinated PI-based frequency deviation control of isolated hybrid microgrid: An online multi-agent tuning approach via reinforcement learning. In *2022 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*, pages 1–5, 2022

5. V. Skiparev, J. Belikov, E. Petlenkov, and Y. Levron. Reinforcement learning based MIMO controller for virtual inertia control in isolated microgrids. In *IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)*, Novi Sad, Serbia, 2022

6. V. Skiparev, K. Nosrati, A. Tepljakov, E. Petlenkov, Y. Levron, J. Belikov, and J. M. Guerrero. Virtual inertia control of an isolated microgrid using NN-VFOPID controller: A self-tuning approach. *IEEE Transactions on Sustainable Energy*, 2023

7. K. Nosrati, V. Skiparev, A. Tepljakov, E. Petlenkov, J. Belikov, and Y. Levron. Constrained intelligent frequency control in an AC microgrid: An online reinforcement learning based PID tuning approach. In *2023 IEEE Power and Energy Society General Meeting*, pages 1–5, 2023

8. V. Skiparev, K. Nosrati, J. Belikov, A. Tepljakov, and E. Petlenkov. An enhanced NN-based load frequency control design of MGs: A fractional order modeling method. In *2023 IEEE International Conference on Compatibility, Power Electronics and Power Engineering (CPE-POWERENG)*, Tallinn, Estonia, 2023