



# Article Management of Voltage Flexibility from Inverter-Based Distributed Generation Using Multi-Agent Reinforcement Learning

Nikita Tomin<sup>1,\*</sup>, Nikolai Voropai<sup>1</sup>, Victor Kurbatsky<sup>1</sup> and Christian Rehtanz<sup>2</sup>

- <sup>1</sup> Melentiev Energy Systems Institute SB RAS, Elecric Power Systems Department, 664033 Irkutsk, Russia; voropai@isem.irk.ru (N.V.); kurbatsky@isem.irk.ru (V.K.)
- <sup>2</sup> Institute of Energy Systems, Energy Efficiency and Energy Economics (ie3), TU Dortmund University, 44227 Dortmund, Germany; christian.rehtanz@tu-dortmund.de
- Correspondence: tomin.nv@gmail.com

**Abstract**: The increase in the use of converter-interfaced generators (CIGs) in today's electrical grids will require these generators both to supply power and participate in voltage control and provision of grid stability. At the same time, new possibilities of secondary QU droop control in power grids with a large proportion of CIGs (PV panels, wind generators, micro-turbines, fuel cells, and others) open new ways for DSO to increase energy flexibility and maximize hosting capacity. This study extends the existing secondary QU droop control models to enhance the efficiency of CIG integration into electrical networks. The paper presents an approach to decentralized control of secondary voltage through converters based on a multi-agent reinforcement learning (MARL) algorithm. A procedure is also proposed for analyzing hosting capacity and voltage flexibility in a power grid in terms of secondary voltage control. The effectiveness of the proposed static MARL control is demonstrated by the example of a modified IEEE 34-bus test feeder containing CIGs. Experiments have shown that the decentralized approach at issue is effective in stabilizing nodal voltage and preventing overcurrent in lines under various heavy load conditions often caused by active power injections from CIGs themselves and power exchange processes within the TSO/DSO market interaction.

**Keywords:** voltage flexibility; droop control; multiagent reinforcement learning; hosting capacity; active distribution system; microgrid

# 1. Introduction

Maintaining busbar voltage within specified levels throughout the entire power system is essential for stability and power quality. Voltage stability is largely related to the reactive power capabilities of generating units and reactive power compensators and their location. Demand behavior and the presence of distributed energy resources (DERs) have a considerable impact on voltage. In this case, an increase in the number of DERs creates completely new power flows, alters the voltage profiles in the distribution system, and can diminish power quality. In this context, there is an elevated need for flexible solutions to maintain the required voltage level at busbars. Ancillary services from distributed inverter-based generation and energy storage systems can be instrumental solutions for voltage stabilization in systems with DERs [1].

It is worth noting that, for example, European draft network codes prohibit or restrict the *DSO*'s ability to export reactive power to transmission networks. There are also physical operating limits in terms of voltage to be met [1]. As the number of distributed energy resources connected to distribution networks increases, active power injections lead to changes in voltage profile. However, with effective coordination, *DSO*s can employ converter-interfaced generation (CIG) to monitor voltage and manage power losses. Moreover, *DSO*s, by using flexibility services, could better control voltage profiles in areas with a large number of renewable energy sources (RESs).



Citation: Tomin, N.; Voropai, N.; Kurbatsky, V.; Rehtanz, C. Management of Voltage Flexibility from Inverter-Based Distributed Generation Using Multi-Agent Reinforcement Learning. *Energies* 2021, 14, 8270. https://doi.org/ 10.3390/en14248270

Academic Editor: Ahmed Abu-Siada

Received: 12 October 2021 Accepted: 6 December 2021 Published: 8 December 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). Thus, flexibility can directly benefit load-controlled consumers (for example, solar panel owners) that can supply more energy to the grid. It also means an increase in hosting capacity, which is understood as the maximum DERs capacity when connected to the grid, when the voltage at all nodes is within acceptable limits, and the current flowing through the network elements, including the transformer and lines, is within limits indicated in [2]. The value of increased flexibility and hosting capacity is also driven by deferred investment and reduced costs of maintaining voltage.

As a result, the need for flexibility, classified as voltage flexibility, is necessary to stabilize it locally and regionally. The time scale related to voltage flexibility ranges from seconds to tens of minutes [1].

### 1.1. Problem Statement

New principles for adaptive CIG management are needed to maximize flexibility for various local and system services. At the same time, the integration of distributed generation (DG) into distribution networks (Figure 1) brings about the following technical problems:



Figure 1. Converter-interfaced generation (CIG) integration into the existing electrical networks.

- overvoltage at nodes where DG "delivers" significant amount of active power to the network;
- overload of distribution transformers and lines;
- other voltage problems such as imbalance, power quality problems (flicker, voltage wave quality);
- mis-operation of relay protection systems due to bidirectional power flows.

In the light of CIG integration, the development of microgrids, which can be considered as localized distribution sub-networks, is a particular issue as they perform independent control, including that after disconnection from the main network through the point of common coupling (PCC) (Figure 1). At the same time, in urban distribution networks, which can include several microgrids, in the feeders of a network in a residential area, where the load is minimal but the amount of solar energy is significant during the daytime, solar generation can cause a reverse energy flow into the network. Due to the relatively high resistance of the distribution network, such a reverse power flow can cause overvoltage limiting the use of existing renewable energy sources and the integration of new photovoltaic plants. All these factors determine the problem of maximizing hosting capacity. In this regard, the primary purpose of calculating PV hosting capacity is to inform energy suppliers about the limitations of the possibility of integrating PV modules for their feeder without the need to upgrade the network [3]. Consequently, given the recent advances in voltage control techniques and more extensive deployment of CIGs, and, above all, PV modules, it is necessary to include voltage control in the analysis of the placement of such modules.

### 1.2. Related Work

Classical control methods in such cases may have limitations, such as compensation for reactive power that can lead to overloads. This factor, in particular, affects the fact that the existing methods for placing PV modules do not factor in voltage control devices when calculating PV hosting capacity [4,5]. This is because outdated voltage monitoring devices, such as LTCs and capacitor banks, are not fast enough, which is why transient overvoltage occurs. At the same time, modern intelligent inverters can maintain the reactive power of the feeder when operating in several control modes. They are faster than traditional controllers and are, therefore, potential candidates for eliminating voltage quality problems arising from the variability of DERs.

According to [6–10], provision of optimal operating conditions and a further improvement in CIG performance, for example, maximization of hosting capacity in distribution networks, reduction of network losses, and others, require enhanced state estimation and coordinated operation of various control means. At the same time, even a low level of communication between CIGs allows achieving better control settings and increased performance [11]. In the future, in addition to P- and Q-control-based flexibility services, CIGs may also provide other local power quality improvement services for DSOs.

One of the promising solutions to the above problems can be secondary voltage control through the coordination of available QU-droop control-based regulators. A new approach to integrating CIGs into distribution grids and micro-grids suggests harmonized voltage control using inverters, through which solar and/or wind generation is connected to the grid and which are located at the end users. In the context of considerable DER-based generation and a plunge in consumption from the grid, voltage stabilization and loss reduction are provided by remotely controlled inverters standing "behind the meter". Various flexibility services related to active power P and reactive power Q from CIG units can be provided by different modes of primary control of the inverter, through which different types of DERs are connected [12]. More specifically (Figure 2), the primary controller of each DG i, i = 1, ..., N registers reference voltages,  $V_{ni}$ , from the secondary controller and regulates output voltage  $V_{oi}$  to the required setting, which is usually achieved using reactive droop control (Q/V strategy) methods without data exchange between CIGs [13,14].



Figure 2. Diagram of inverter-based distributed generation (DG).

Existing methods of secondary control can be divided into two main classes: centralized and distributed. The centralized controller collects information from all CIGs and makes a decision on collective management of the electrical network operation, which is then sent to the appropriate CIGs. A vivid practical example of such a management is that of the California Independent System Operator (CAISO), which is considered to be a bold experiment on energy flexibility. Based on the 131 MW Tule wind farm in San Diego, CAISO and Avangrid Renewables have proved that powerful wind turbines connected to the grid through inverters are sources of energy flexibility and can successfully provide services to control frequency and active power flows, and maintain voltage [15]. In the proposed solution, a centralized PCC controller is responsible for the function of critical control of all inverters in a wind farm, and it continuously monitors the state of inverters and controls them to ensure that they produce the active and reactive power needed to provide the desired voltage curve on the high side of the transformer.

Although centralized methods of control for CIG-based systems show promising results, such methods are associated with loss of bandwidth in the communication lines and are often associated with the issues of a single point of failure, as well as the "curse of dimension", which makes it impractical to deploy them in today's large power systems [16]. Alternatively, one can employ distributed methods, where each CIG interacts with neighboring CIGs and decides on decentralized control based on its state and the states of its neighbors shared through local communication networks.

At the same time, the basic principle of such approaches is the exchange of information through neighboring communication using a distributed protocol and reaching a consensus, for example, an average value of the measured voltages. In contrast to frequency, voltages are local variables, which means that they can be restored either on selected critical buses or at a system level [16]. In the latter case, the distributed methods can be used to generate a common signal, which is compared with a reference one and passes through a local PI controller that generates an appropriate control signal to be sent to the primary level to eliminate associated steady-state errors. Traditional distributed secondary controllers were based on the principle of normal averaging [17,18]. These papers defined the interaction between CIGs as a key component in achieving the control aims while avoiding a centralized architecture. The published works also present several distributed control methods, of which the algorithms of gossip [19] and consensus [20] have recently drawn considerable attention, mainly due to their robustness for distributed information exchange over networks. Given the specific features of the distributed nature of control, decentralized approaches often use a multi-agent systems (MAS) framework [21].

Thus, traditional principles of distribution network operation and control limit the CIG's capabilities to provide system-wide ancillary services in certain situations. Overcoming these limitations requires new principles of active and adaptive control [2]. Therefore, with recent advances in voltage control methods and large-scale deployment of DERs with intelligent inverters, it is necessary to include voltage control in the analysis of CIG placement. With this approach, it is possible to simultaneously achieve better hosting capacity of the distribution network and the flexibility of CIG services, even in very low-load situations.

### 1.3. Paper Contribution

The aim of this paper is to extend the existing multi-agent systems (MAS) models of decentralized inverter-based secondary voltage control to improve CIG-associated integration problems (overvoltages, voltage flexibility, and hosting capacity) in active distribution networks and microgrids. The paper proposes a new approach to the decentralized inverterbased secondary voltage control based on multi-agent deep reinforcement learning (MARL) algorithm to improve voltage flexibility and hosting capacity of microgrids and active distribution networks. The proposed approach can help better maintain voltage, maximize hosting capacity in distribution networks, and improve the availability of distribution network-connected DERs for TSO flexibility services. We adopt the centralized training and decentralized execution scheme, where each agent has its actor and critic networks, and their policies are updated independently in contrast to the algorithm of consensus that may hurt the convergence speed.

The remainder of the paper is organized as follows. Section 2 describes the proposed methodological MARL-based distributed voltage droop control framework as well as estimating hosting capacity/voltage flexibility approach. Section 3 presents a case study based on a modified MV IEEE 34-bus test feeder to demonstrate the main features of the

MARL model. Section 4 summarizes the main findings and highlights the ideas for further research.

### 2. Materials and Methods

### 2.1. Voltage Droop Control for Inverters

Inspired by droop control used for synchronous generators, researchers have proposed a similar control scheme to inverters [21–24]. The primary motivation for this is that droop control actually implements decentralized proportional control and, therefore, represents a plug-and-play-like control scheme that is modular and hence simple in implementation in the sense that there is no need for centrally coordinated network control. In large highvoltage transmission systems, droop control is usually used only to obtain the desired active power distribution, while the voltage amplitude on the generator bus is regulated to the nominal voltage setpoint using (usually in the range of  $0.95 \div 1.05$  p.u.) a power system stabilizer. However, unlike high voltage transmission systems (hundreds–thousands of kilometers), the transmission lines in microgrids are usually relatively short (few tens of kilometers), which is why droop control is employed here to control voltage to achieve the desired reactive power distribution.

The rationale for using voltage droop controllers is as follows [25]. It follows for small angular deviations  $\delta_{ik}$ , that  $\sin \delta_{ik} \approx \delta_{ik}$ , and  $\cos \delta_{ik} \approx 1$ . Consequently, reactive power in predominantly inductive networks, i.e., where  $G_{ik} \approx 0$ , is most affected by voltage changes. Therefore, amplitudes of the invertor voltage  $V_i$  vary depending on reactive power deviations (in terms of the desired value) according to:

$$u_i^V = V_i^d - k_{Qi} \left( Q_i^m - Q_i^d \right) \tag{1}$$

where  $V_i^d \in \mathbb{R} > 0$  is desired (nominal) voltage amplitude,  $k_{Qi} \in \mathbb{R}^+$  is voltage gain,  $Q_i^m : R \ge 0 \rightarrow R$  is measured reactive power, and  $Q_i^d \in R$  is its desired settings.

For predominantly inductive networks and small angular deviations (for instance microgrids in islanded mode with sudden switching of reactive load), reactive power flow of the *i*-th node  $Q_i$ ,

$$Q_i(\delta_1,\ldots,\delta_n,V_1,\ldots,V_n) = G_{ii}V_i^2 + \sum_{k\sim\mathcal{N}_i} |Y_{ik}|V_iV_k\sin(\delta_{ik}+\phi_{ik})$$
(2)

decreases to  $Q_i$ :  $\mathbb{R}^n_{>0} \to \mathbb{R}$ :

$$Q_i(V_1,\ldots,V_n) = |B_{ii}|V_i^2 + \sum_{k \sim \mathcal{N}_i} |B_{ik}|V_iV_k$$
(3)

In this case, then, reactive power  $Q_i$  can be controlled by controlling amplitudes of voltage  $V_i$  and  $V_k$ ,  $k \sim N_i$ .

# 2.2. Multi-Agent Reinforcement Learning (MARL)-Based Distributed Voltage Control for Inverters

Reinforcement learning is one of the machine learning methods, during which the system (agent) under test learns by interacting with some environment. Reinforcement signals are the response of the environment to decisions made. The environment is usually formulated as a Markov decision-making process with a finite set of states. Formally, the simplest reinforcement learning model consists of a set of environmental states *S*, a set of actions *A*, and a set of scalar "gains". At any time instant *t*, agent is characterized by state  $s_t \in S$  and set of potential actions  $a \in A(s_t)$ ; it transitions to state  $s_{t+1}$  and gains a reward  $r_t$ . Based on this interaction with the environment, the reinforcement learning agent must strategize,  $\pi : S \times A \rightarrow [0, 1]$ , where  $\pi(s, a)$  is the probability of choosing an action  $a \in A(s_t)$  in state *s*. This strategy maximizes the value  $R = r_0 + r_1 + \cdots + r_n$  in the Markov decision-making process [26].

error algorithm.

MARL is an extension of the single-agent model and refers to multi-agent/player systems. In recent years, several MARL-based approaches have been proposed for autonomous voltage control in microgrids [27–30]. The agent can find optimal policies, when they interact with the environment as well as offline learn to cooperate with other agents by simulating their policies. After completing training, agents can make real-time decisions that adapt well to unknown power grid or microgrid dynamics. This fact determines a strong motivation of developing MARL-based voltage control applications for isolated microgrids and energy communities with RESs and power flexibility services. Based on the analysis of these works, in this paper, we have developed a MARL-based model-free approach for decentralized inverter-based secondary voltage control to manage flexibility services and increase hosting capacity. By model-free algorithms are meant that do not actually use the well-known model of the environment associated with the Markov decision process. In fact, this type of reinforcement learning method can be thought of as a trial and

Multi-agent networks can be represented as graphs in which vertices represent physical or virtual items (agents) and edges represent the interaction between them. Specifically, we model the electrical network with CIG as a multi-agent network,  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ , where each agent  $i \in \mathcal{V}$  interacts with its neighbors  $N_i : \{j | \varepsilon_{ij} \in \mathcal{E}\}$ . Then we can consider *S* and  $\mathcal{A}$  as the global state and action spaces that represent, respectively, aggregated set on state and control for all CIGs. The main dynamics of the microgrid can present using the state transition probability  $\mathcal{P} : S \times \mathcal{A} \rightarrow [0, 1]$ . We consider a decentralized MARL framework to achieve scalable inverter-based secondary voltage control. Each CIG only communicates with its neighbors and makes control decisions based on these observations. Since each agent i (CIG i) observes only part of the environment (its own state and the state of its neighbors), we have a partially observable Markov decision process (POMDP) [31].

We solve the above problem with MARL and define the key elements in the POMDP in question as follows:

- Action space: the control action for each CIG is the secondary voltage control setpoint  $V_n$ . By analogy with [30], we used 10 discrete actions evenly distributed between 1.00 and 1.14 p.u. The overall action of a microgrid or active distribution network is the joint actions of all DG, i.e.,  $a = v_{n1} \times v_{n2} \times \cdots \times v_{nN}$ .
- State space: the state of each CIG i is chosen as  $s_t = (\delta_i, P_i, Q_i, i_{odi}, i_{oqi}, i_{bdi}, i_{bqi}, v_{bdi}, v_{bqi})$ to characterize operating parameters of CIGs, where  $\delta_i$  is measured reference angle (phase);  $P_i, Q_i$  are active power and reactive power, respectively;  $i_{odi}, i_{oqi}, i_{bdi}, i_{bqi}$  [A] are output currents d-q of CIG i and directly connected busbars, respectively; while  $v_{bdi}, v_{bqi}$  [kV] are output voltages d-q of the connected busbar, respectively.
- Space of observations: it is assumed that each CIG can only observe its local state and messages from its neighbors, i.e.,  $o_{i,t} = S_{i,t} \cup m_{i,t}$ , where  $m_{i,t}$  is communication message received from neighboring agents  $j \in N_i$ , which will be considered further in more detail.
- Transition Probabilities: the probability of transition  $T(s_0|s, a)$  is a characteristic of the dynamics of the electrical network with CIG. We follow the models from [32] to build a platform for simulating the operating conditions of a microgrid or active distribution network without using any prior knowledge of the transition probability since the MARL used is model-free.
- Reward function: we apply the following reward function for generators to converge quickly to reference voltages (for example, one p.u.):

$$r_{i,t} = \begin{cases} 0.05 - |1 - v_i|, & v_i \in |0.95, 1.05|, \\ -|1 - v_i|, & v_i \in |0.8, 0.95| \cup |1.05, 1.25| \\ -10. & Otherwise \end{cases}$$
(4)

where  $r_{i,t}$  is a reward of agent *i* at time step *t*. We split the voltage range into three working areas similarly to [30]. These are an area of normal operating conditions (|0.95, 1.05| p.u.),

an area of heavy load conditions ( $|0.8, 0.95| \cup |1.05, 1.25|$  p.u.), and an emergency area ( $|0, 0.8| \cup |1.25, \infty|$  p.u.). With the reward formulated, CIGs with "emergency" voltages will receive a high penalty, while CIGs with voltages close to 1 p.u. will receive a positive reward.

The proposed voltage control is distributed and requires communication among CIGs in the network. We consider a decentralized MARL structure in which each agent (CIG) can communicate with its neighbors and exchange necessary information, for example, states. Information from neighboring agents is used to enhance the efficiency of training. Thus, based on the structure proposed in [30], agent *i* updates its hidden state  $h_{i,t}$  at each step *t*.

$$h_{i,t} = f_i(h_{i,t-1}, q_0(e_s(o_{i,t})), q_h(h_{\mathcal{N},t-1}))$$
(5)

where  $h_{i,t-1}$  is a hidden state from the previous time step;  $o_{i,t}$  is the observation of agent i, which was made at time t, i.e., its internal state and the states of its neighbors;  $h_{N,t-1}$  is an integrated state from neighbors;  $e_s$ ,  $q_0$ , and  $q_h$  are differentiable message encoding and extraction functions that use single-layer fully connected deep neural network layers with 64 neurons; while  $f_i$  is the function of encoding hidden states and communication information, where we use the LSTM network. In this article the deep neural structure was chosen based on the studies obtained in [31], where the authors introduced the deep recurrent Q-network (DRQN), a combination of a LSTM and a Deep Q-Network. Such approach shown better results to solve POMDPs than comparable (non-LSTM) neural networks.

Instead of low-dimensional indicators, as in [33], we include the neighbor's complete states in the local observation  $o_{i,t} = s_{i,t} \cup s_{N,t}$ , to improve the observability of the agent and use the network to automatically examine the corresponding representation. In this case, the received communication message  $m_{i,t}$  of the *i*-th agent is a combination of internal states and hidden states of its neighbors.

Hidden state  $h_{i,t}$  received from (5) is then used in actor-critic networks to generate random actions and predict value functions, respectively, i.e.,  $\pi_{\theta_i}(|h_{i,t}|)$  and  $V_{\omega_i}(h_{i,t})$  (Figure 3). We use a centralized training scheme with decentralized execution [34,35], where each agent has its actor-critic networks, and their policy is updated independently but not based on consensus [36] that can reduce the convergence rate of the solution.





**Figure 3.** A general diagram of the proposed approach for inverter-based secondary voltage control using the multi-agent reinforcement learning (MARL) structure.

Cooperative MARL aims to maximize total global rewards  $R_{g,t} = \sum_{i \in \mathcal{V}} R_{i,t}$ , where  $R_{i,t} = \sum_{k=0}^{T} \gamma^k r_{i,t+k}$  denotes the cumulative reward for agent *i*. Such a mathematical formulation, however, is associated with typical problems of multi-agent training [26]. These are loss of bandwidth, possible decrease in training efficiency, restrictions on the number of agents, and slow convergence of the global solution. The spatial discounting factor was

proposed in [30] to solve these problems when each agent *i* uses the following reward:  $R_{i,t} = \sum_{k=1}^{T} \gamma^{k} \sum \alpha(d_{i,i}) r_{i,t+k}$ (6)

$$R_{i,t} = \sum_{k=0}^{1} \gamma^k \sum_{j \in v} \alpha(d_{i,j}) r_{i,t+k}$$
(6)

where  $\alpha(d_{i,j}) \in [0, 1]$  is a spatial discounting function,  $d_{i,j}$  is a distance between agents *i* and *j*. The distance can be the Euclidean distance, which characterizes the physical distance between two agents (generators), or the distance between two vertices on the graph (i.e., the number of shortest connecting edges).

### 2.3. Estimating Photovoltaic (PV) Hosting Capacity and Voltage Flexibility

MV and LV networks have limited hosting capacity, which depends on load conditions, the capacity of components, and network topology. Excess of this limit manifests itself through overvoltage, undervoltage (voltage limitation), or line or transformer overload (current limitation). In networks with voltage-limited hosting capacity, intelligent inverters can provide additional network flexibility and increase hosting capacity [37].

Although hosting capacity now generally refers to various types of CIGs, in this paper, we focus on the classical PV hosting capacity analysis to assess the performance of a decentralized MARL-based inverter control method. The basic idea of PV hosting capacity calculation, in this case, is to increase the number of PV plants in the distribution network or microgrids until any scheduling principle or limitation is violated. We assess the feeder's PV hosting capacity, which means the largest capacity of a PV plant that can be placed without violating operational restrictions. In this case, we focus on overvoltage and overload in lines and MV-LV transformers.

For the stochastic nature of PV module placement to be factored in, we use the Monte Carlo simulation approach to simulate a whole host of different future PV installation scenarios [4]. We modify the algorithm proposed in [37] by simulating k scenarios of placing PV modules, each of which represents one Monte Carlo run for the investigated distribution network or microgrid. Then PV hosting capacity, H, can be defined as:

$$H = \min_{i \in S} \left\{ PV_{pen}^{i} \left| P\left(V_{max,k}^{i} > 1.05; \ I_{Tmax,k}^{i} > 50\%; \ I_{Lmax,k}^{i} > 50\%\right) = 1 \right\}$$
(7)

where *S* is the discrete *PV* customer penetration levels, indexed by  $i, S \in \{1, 2, ..., i, ..., 100\}$ ;  $PV_{pen}^{i}$  is the set of all *PV* penetration levels indexed by customer penetration level *i*,  $\{PV_{pen}^{1}, PV_{pen}^{2}, ..., PV_{pen}^{i}, ..., PV_{pen}^{100}\}$ ;  $V_{max,k}^{i}$ ,  $I_{Tmax,k}^{i}$ ,  $I_{Lmax,k}^{i}$  is the set of maximum primary voltages, line loading and transformer loading recorded for *k PV* deployment scenarios.

Based on the aforementioned, to relate the concepts of flexibility and hosting capacity for active distribution networks, we numerically estimate the voltage flexibility as:

$$FLEX_V = \frac{H_{base} - H_{cont}}{H_{base}} 100\%$$
(8)

where  $H_{base}$  is *PV* hosting capacity calculated for the base case without voltage regulation;  $H_{cont}$  is *PV* hosting capacity calculated for the option with the possibility of voltage control. It is worth noting that  $H_{cont}$  suggests calculation of *PV* hosting capacity with control instruments available in the distribution network or microgrid, including QU-droop control, regulation of the transformer tap, control of compensating devices, and others.

# 3. Results

We applied the proposed MARL-based decentralized voltage control approach to a modified MV IEEE 34-bus test feeder with six CIGs. This system was an actual feeder located in Arizona. Its nominal voltage is 24.9 kV. It is characterized by long and lightly loaded two in-line controllers, an in-line transformer for a short 4.16 kV section, unbalanced load, and shunt capacitors. This system was designed to evaluate and benchmark algorithms in solving unbalanced three-phase radial systems. Thus, this system represents a reduced-order model of an actual distribution circuit. In our modification, this network includes six CIGs (PV systems, each 20 kW) and a somewhat simplified topology sufficient to demonstrate the proposed secondary QU control approach (Figure 4). The main parameters of CIGs, lines, and loads are summarized in Table 1.



Figure 4. A modified IEEE 34-bus test feeder with six inverted-based DGs.

	DG1, DG2, DG5, DG6		DG3, DG4	
	$m_p$	$5.64 imes10^{-5}$	$m_p$	$7.5 imes10^{-5}$
	$n_O$	$0.52  imes 10^{-3}$	$n_O$	$0.60 imes10^{-3}$
CIG	$R_c^{\sim}$	0.03 Ω	$\widetilde{R_c}$	0.03 Ω
	$L_c$	0.35 mH	$L_c$	0.35 mH
	$w_c$	31.41	$w_c$	31.41
	$k_p$	4	$k_p$	4
	$\dot{k_i}$	40	$k_i$	40
	Load 1	Load 2	Load 3	Load 4
Loads	1.5 Ω	0.5 Ω	1 Ω	0.8 Ω
	0.03 Ω	0.017 Ω	0.05 Ω	0.02 Ω

Table 1. Specifications of the IEEE 34-bus feeder system.

#### 3.1. MARL-Based QU Droop Control

The MARL approach is implemented in the Python environment using open-source tools for power system modeling (pandapower and PowerNet). The simulation platform used is based on the technical characteristics of line and load described in [35,38]. Simulation of heavy load conditions involved random load changes added throughout the network with deviations of  $\pm 20\%$  from the nominal values and random disturbances in the range of  $\pm 5\%$  for each load. All CIGs in the considered schemes were monitored with

a sampling time of 0.05 s, and each CIG could communicate with its neighbors across local boundaries of communication. The primary control of the lower level is implemented by an analogy with [32].

We compare the used MARL approach with several state-of-the-art benchmark MARL algorithms: IA2L [39] and CommNet [40], to demonstrate its effectiveness. We train each model over 10,000 episodes, with  $\gamma = 0.99$ , minibatch size N = 20, actor learning rate  $\eta_{\theta} = 5 \times 10^{-4}$ , and critic learning rate  $\eta_{\omega} = 2.5 \times 10^{-4}$ . To ensure fair comparison, each episode generates different random seeds and in each episode the same random seed is shared across different algorithms to guarantee the same training/testing environment. We control the agents every (a simulation time)  $\Delta T = 0.05$  s and one episode lasts for T = 20 steps.

Figure 5a shows the training curve of the MARL algorithm for the modified IEEE 34-bus feeder. It is clear that the used MARL outperforms these state-of-the-art MARL algorithms in terms of convergence speed. After 5000 training episodes, the obtained strategy was assessed 20 times for various load disruptions with the same random seed for each agent in each episode.



**Figure 5.** Performance trained MARL policy under 25% load disturbance: (a) MARL training curves for IEEE 34-bus feeder system. The lines show the average reward per training episode, which is smoothed over the past 100 episodes; (b) performance of voltage control.

The results of this testing are presented in Table 2 and Figure 5b that show the voltage profiles for nodes with inverter generators for simulation of one of the heavy load conditions of the system (load increase by 25%). As noted above, the secondary QU control aims to bring all DGs voltages to a reference value of 1 p.u. As seen in Figure 5b, in the case of a voltage drop, MARL-control in 0.4 s after the disturbance starts restores voltage to its nominal values.

Table 2. Performance trained MARL policies under different load disturbances.

Load Disturbance	Average Reward over 20 Evaluation Episodes
5%	0.27
10%	0.24
15%	0.23
25%	0.22

Additionally, the effects of decentralized voltage control are demonstrated by representing the results of operating parameters calculation on the IEEE 34-bus feeder graph for various experimental cases (Figure 6). Comparison of Figure 6b,c indicates that the secondary QU control not only stabilizes voltage at nodes but also reduces overcurrent in system lines. For example, under heavy load conditions, the overloads in Line 0 and Line 1 are 122.68% and 44.36%, respectively (Figure 6b). Secondary QU control leads to a decrease in the current overload in these lines to 91.08% and 32.83% (Figure 6c), respectively.



**Figure 6.** Calculated operating parameters plotted on the graph of the modified IEEE 34-bus feeder for various simulation experiments. (a) Normal operating condition; (b) heavy load condition (non-control); (c) heavy load condition (secondary QU control).

### 3.2. Voltage Flexibility and Hosting Capacity Analysis

An analysis of PV hosting capacity, according to (7), relied on the Monte Carlo method employed for probabilistic modeling of a large number of various scenarios for the installation of future PV plants in the IEEE 34-bus feeder. Each scenario consisted of PV systems of a certain capacity connected to specific nodes of the system. The obtained statistical distribution of the maximum installed capacity of PV systems is an additional hosting capacity of the network. The assessment of the effect of the secondary QU control on hosting capacity rests on two experiments: with and without QU control of inverters. An approximate distribution of the maximum number of installed PV plants for these two experiments is shown in Figure 7. The results were obtained by simulating 50 different possible scenarios for the future installation of PV modules.

The results for the scenario without voltage control show that the hosting capacity of the IEEE 34-bus feeder ranges from 12.8 MW to 17.6 MW overall (Figure 7a). We can also note that for 50% of the runs it is between 14.1 and 15.3 MW (the median equal 4.8 MW of additional PV capacity). The results also show that the potential problems due to connection of additional PV plants arise due overloading of a transformer (in 86% of cases), and a violation of the voltage band (in 14% of the cases). In the scenario when we have a voltage droop control, the hosting capacity ranges from 13.3 MW to 17.9 MW overall, and the median is increased to about 15.3 MW (Figure 7b). As result, the figure shows that with intelligent MARL-control of inverters, the hosting capacity of the considered electrical network increases from H = 14.6 MW to H = 15.7 MW. The box plot showing the resulting distribution helps to understand the behavior of the network when more PV systems are

installed. It shows the minimum, maximum, and average number of additionally installed PV systems that are the first to expect violations. As a result, with intelligent control of inverters, these potential violations (when the maximum hosting capacity is exceeded) are reduced only to an overload of transformers (which we do not control), but voltage violations was eliminated (Figure 7b).



**Figure 7.** Results of the hosting capacity analysis for a modified IEEE 34-bus feeder for various experiments. (**a**) Non-control case; (**b**) secondary QU control.

Expression (8) was used to evaluate the effect of voltage flexibility rise when using MARL- control of inverters. For the considered series of experiments with a modified IEEE 34-bus feeder,  $FLEX_V = 7.53$  %. This means that the obtained difference  $H_{base} - H_{cont} = 1.1$  MW determines additional power that can be used to optimize operating conditions of the distribution network (already within the framework of tertiary control), for example, when selling electricity at the TSO level, without the risk of voltage problems and possible overcurrent in lines.

# 4. Conclusions

The future will require active use of flexible energy resources connected to the distribution network, including those connected to the low voltage 0.4 kV grid, to provide flexible DSO and TSO services across new markets. New possible instruments of secondary control in distribution networks and microgrids with a large share of CIG, including the coordination of adaptive droop Q/U-controllers will further increase operational flexibility, hosting capacity, and degrees of freedom in TSO/DSO interaction within the framework of market interaction.

We have proposed an approach to decentralized secondary voltage control through inverter generation based on the MARL algorithm to assess new opportunities for increasing voltage flexibility in distribution networks with a considerable proportion of CIGs. With this approach, the electrical network is considered as a multi-agent one, where each agent (CIG) learns a control policy based on (sub-) global reward, local states, and encoded communication messages from its neighbors (other CIGs). Experimental studies based on a modified IEEE 34-bus feeder have shown that this approach can effectively stabilize the voltage at network nodes and prevent overcurrent in lines under heavy load conditions. In the context of TSO/DSO market interaction, such operating conditions can result from the power exchange between systems of different voltage levels during flexibility services rendered. Findings indicate that the proposed approach to secondary QU-control allows increasing the voltage flexibility of the network at the DSO level and maximizing the hosting capacity.

**Author Contributions:** Conceptualization, N.T. and C.R.; Data curation, N.T.; Formal analysis, N.T., N.V.; Funding acquisition, N.T.; Investigation, N.T.; Methodology, N.T., V.K.; Project administration, N.T.; Software, N.T.; Supervision, C.R.; Validation, N.T.; Visualization, N.T.; Writing—original draft, N.T., V.K. and N.V.; Writing—review & editing, N.T. All authors have read and agreed to the published version of the manuscript.

**Funding:** The reported study was funded by the Russian Science Foundation (No. 19-49-04108) and the German Science Foundation/DFG (No. RE 2930/24)—Sections 1, 2.1, 2.2 and 4.1; the Russian Foundation for Basic Research (RFBR) No. 21-58-53049— Section 2.3 and 4.2.

Conflicts of Interest: The authors declare that there is no conflict of interest.

### References

- Hillberg, E.; Zegers, A.; Herndler, B.; Wong, S.; Pompee, J.; Bourmaud, J.Y.; Lehnhoff, S.; Migliavacca, G.; Uhlen, K.; Oleinikova, I.; et al. Flexibility Needs in the Future Power System. ISGAN Annex 6 Power T&D Systems. 2019. Available online: https://www.iea-isgan.org/wp-content/uploads/2019/03/ISGAN\_DiscussionPaper\_Flexibility\_Needs\_In\_Future\_Power\_ Systems\_2019.pdf (accessed on 12 August 2021).
- 2. Laaksonen, H.; Parthasarathy, C.; Hafezi, H.; Shafie-khah, M.; Khajeh, H.; Hatziargyriou, N. Solutions to Increase PV Hosting Capacity and Provision of Services from Flexible Energy Resources. *Appl. Sci.* **2020**, *10*, 5146. [CrossRef]
- 3. Abad, M.S.S.; Ma, J. Photovoltaic Hosting Capacity Sensitivity to Active Distribution Network Management. *IEEE Trans. Power Syst.* **2021**, *36*, 107–117. [CrossRef]
- 4. Smith, J. Stochastic Analysis to Determine Feeder Hosting Capacity for Distributed Solar PV; EPRI Technical Report 1026640; EPRI 3420 Hillview Avenue: Palo Alto, CA, USA, 2012.
- Dong, Y.; Wang, S.; Yu, L. Voltage Sensitivity Analysis Based PV Hosting Capacity Evaluation Considering Uncertainties. In 2020 IEEE Power & Energy Society General Meeting (PESGM); IEEE: Piscataway, NJ, USA, 2020; pp. 1–5.
- 6. Ali, W.; Ulasyar, A.; Mehmood, M.U.; Khattak, A.; Imran, K.; Zad, H.S.; Nisar, S. Hierarchical Control of Microgrid Using IoT and Machine Learning Based Islanding Detection. *IEEE Access* **2021**, *9*, 103019–103031. [CrossRef]
- Muttaqi, T.; Baldwin, T.L.; Chiu, S.C. Distribution System State Estimation with AMI Based on Load Correction Method. In 2019 North American Power Symposium (NAPS); IEEE: Piscataway, NJ, USA, 2019; pp. 1–6.
- 8. Alahäivälä, A.; Saarijärvi, E.; Lehtonen, M. Modeling Electric Vehicle Charging Flexibility for the Maintaining of Power Balance. *Int. Rev. Electr. Eng.* **2013**, *8*, 1759–1770.
- 9. Tomin, N.; Maass, J.; Domyshev, A. Flexible Charging Optimization for Electric Vehicles using MDPs-based Online Algorithms. *IFAC-PapersOnLine* 2020, *53*, 12614–12619. [CrossRef]
- 10. Wang, H.; Yan, Z.; Shahidehpour, M.; Zhou, Q.; Xu, X. Optimal Energy Storage Allocation for Mitigating the Unbalance in Active Distribution Network via Uncertainty Quantification. *IEEE Trans. Sustain. Energy* **2021**, *12*, 303–313. [CrossRef]
- Adiguno, F.K.; Mai, T.T.; Nguyen, P.H. Mitigating Impact of Large-Scale PV Integration on MV Distribution Network with Sequential Control Functions: A Case Study in Noordwolde Grid, The Netherlands. In Proceedings of the 25th InternationalConference on Electricity Distribution CIRED 2019, Madrid, Spain, 3–6 June 2019.
- 12. Laaksonen, H.; Khajeh, H.; Parthasarathy, C.; Shafie-khah, M.; Hatziargyriou, N. Towards Flexible Distribution Systems: Future Adaptive Management Schemes. *Appl. Sci.* **2021**, *11*, 3709. [CrossRef]
- 13. Joseph, A.; Smedley, K.; Mehraeen, K. Secure Power Distribution against Reactive Power Control Malfunction in DER Units. *IEEE Trans. Power Deliv.* **2021**, *36*, 1552–1561. [CrossRef]
- 14. Hafezi, H.; Laaksonen, H. Autonomous Soft Open Point Control for Active Distribution Network Voltage Level Management. In Proceedings of the 13th IEEE PowerTech 2019, Milan, Italy, 23–27 June 2019.
- 15. California's Wind Market Has All but Died Out. Could Grid Services Revenue Help? Available online: https://www.greentechmedia.com/articles/read/justin-california (accessed on 2 December 2021).
- 16. 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.
- Tian, X.; Wang, Y.; Wang, F.; Guo, Z.; Dong, Y. An improved droop control strategy for accurate current sharing and DC-BUS voltage compensation in DC microgrid. In Proceedings of the 16th IET International Conference on AC and DC Power Transmission (ACDC 2020), Online, 2–3 July 2020; pp. 1466–1473.

- 18. Ning, B.; Han, Q.-L.; Ding, L. Distributed Finite-Time Secondary Frequency and Voltage Control for Islanded Microgrids With Communication Delays and Switching Topologies. *IEEE Trans. Cybern.* **2021**, *51*, 3988–3999. [CrossRef]
- Aysal, T.C.; Yildiz, M.E.; Sarwate, A.D.; Scaglione, A. Broadcast Gossip algorithms for consensus. *IEEE Trans. Signal Process.* 2009, 57, 2748–2761. [CrossRef]
- 20. Olfati-Saber, R.; Fax, J.A.; Murray, R.M. Consensus and cooperation in networked multi-agent systems. *Proc. IEEE* 2007, 95, 215–233. [CrossRef]
- 21. Katiraei, F.; Iravani, M. Power Management Strategies for a Microgrid with Multiple Distributed Generation Units. *IEEE Trans. Power Syst.* 2006, *21*, 1821–1831. [CrossRef]
- 22. Barklund, E.; Pogaku, N.; Prodanovic, M.; Hernandez-Aramburo, C.; Green, T.C. Energy Management in Autonomous Microgrid Using Stability-Constrained Droop Control of Inverters. *IEEE Trans. Power Electron.* **2008**, *23*, 2346–2352. [CrossRef]
- 23. Iyer, S.; Belur, M.; Chandorkar, M. A Generalized Computational Method to Determine Stability of a Multi-inverter Microgrid. *IEEE Trans. Power Electron.* **2010**, *25*, 2420–2432. [CrossRef]
- Schiffer, J. Stability and Power Sharing in Microgrids. Ph.D. Thesis, Technical University of Berlin, Berlin, Germany, 2017; p. 198.
   Guerrero, J.; Loh, P.; Chandorkar, M.; Lee, T. Advanced Control Architectures for Intelligent MicroGrids—Part I: Decentralized and Hierarchical Control. *IEEE Trans. Ind. Electron.* 2013, 60, 1254–1262. [CrossRef]
- 26. Sutton, R.S.; Barto, A.G. Introduction to Reinforcement Learning; MIT Press: Cambridge, MA, USA, 2018.
- 27. Wang, S.; Duan, J.; Shi, D.; Xu, C.; Li, H.; Diao, R.; Wang, Z. A datadriven multi-agent autonomous voltage control framework using deep reinforcement learning. *IEEE Trans. Power Syst.* 2020, *35*, 4644–4654. [CrossRef]
- 28. Liu, H.; Wu, W. Online multi-agent reinforcement learning for decentralized inverter-based volt-var control. *arXiv* 2020, arXiv:2006.12841. [CrossRef]
- Cao, D.; Hu, W.; Zhao, J.; Huang, Q.; Chen, Z.; Blaabjerg, F. A multi-agent deep reinforcement learning based voltage regulation using coordinated PV inverters. *IEEE Trans. Power Syst.* 2020, 35, 4120–4123. [CrossRef]
- Chen, D.; Li, Z.; Chu, T.; Yao, R.; Qiu, R.; Lin, K. PowerNet: Multi-agent Deep Reinforcement Learning for Scalable Powergrid Control. *arXiv* 2020, arXiv:2011.12354. [CrossRef]
- 31. Hausknecht, M.; Stone, P. Deep recurrent Q-learning for partially observable MDPs. arXiv 2015, arXiv:1507.06527.
- Bidram, A.; Davoudi, A.; Lewis, F.L.; Qu, Z. Secondary control of microgrids based on distributed cooperative control of multi-agent systems. *IET Gener. Transm. Distrib.* 2013, 7, 822–831. [CrossRef]
- 33. Foerster, J.; Nardelli, N.; Farquhar, G.; Afouras, T.; Torr, P.H.; Kohli, P.; Whiteson, S. Stabilising experience replay for deep multi-agent reinforcement learning. *arXiv* 2017, arXiv:1702.08887.
- Chu, T.; Wang, J.; Codecà, L.; Li, Z. Multi-agent deep reinforcement learning for large-scale traffic signal control. *IEEE Trans. Intell. Transp. Syst.* 2019, 21, 1086–1095. [CrossRef]
- 35. Li, T.; Zhang, J.-F. Consensus conditions of multi-agent systems with time-varying topologies and stochastic communication noises. *IEEE Trans. Autom. Control* 2010, 55, 2043–2057. [CrossRef]
- 36. Zhang, K.; Yang, Z.; Liu, H.; Zhang, T.; Basar, T. Fully decentralized multi-agent reinforcement learning with networked agents. *arXiv* **2018**, arXiv:1802.08757.
- Dubey, A. Impacts of Voltage Control Methods on Distribution Circuit's Photovoltaic (PV) Integration Limits. *Inventions* 2017, 2, 28. [CrossRef]
- Thurner, L.; Scheidler, A.; Schäfer, F.; Menke, J.H.; Dollichon, J.; Meier, F.; Meinecke, S.; Braun, M. Pandapower—An Open Source Python Tool for Convenient Modeling, Analysis and Optimization of Electric Power Systems. *IEEE Trans. Power Syst.* 2018, 33, 6510–6521. [CrossRef]
- 39. Sukhbaatar, S.; Fergus, R. Learning multiagent communication with backpropagation. In *Advances in Neural Information Processing* Systems; NYU: Barcelona, Spain, 2016; pp. 2244–2252.
- 40. Lowe, R.; Wu, Y.I.; Tamar, A.; Harb, J.; Abbeel, O.P.; Mordatch, I. Multi-agent actor-critic for mixed cooperative-competitive environments. In *Advances in Neural Information Processing Systems*; NYU: Long Beach, CA, USA, 2017; pp. 6379–6390.