Learning-Based Hierarchical Volt/Var and Demand Flexibility Control in Active Distribution Networks

Mohammad Hashemnezhad and **Petros Aristidou**

Department of Electrical Engineering and Computer Engineering and Informatics, Cyprus University of Technology, Limassol, Cyprus

Abstract: High PV penetration increases voltage variability in Active Distribution Networks (ADNs). While inverter-based Volt/Var Control (VVC) is the primary means of maintaining voltage limits, its effectiveness is constrained once reactive power capability is saturated. To address this, we propose a hierarchical voltage control scheme where Multi-Agent Reinforcement Learning (MARL) coordinates inverter reactive power, and a single aggregator agent adjusts active power from fast flexible loads only when voltage violations persist and reactive headroom is insufficient. An activation gate ensures that flexibility is used sparingly. Case studies on a high-PV feeder show that VVC alone resolves moderate deviations, while conditional flexibility effectively mitigates severe over- and under-voltage with only 10–15% load adjustment.

Keywords: Voltage Regulation; Active Distribution Networks (ADNs); High PV penetration; Volt/Var Control (VVC); Demand flexibility; Multi-Agent Reinforcement Learning (MARL).

1. INTRODUCTION

The rapid growth of photovoltaic generation in Active Distribution Networks (ADNs) is increasing voltage variability, especially in low-voltage feeders with limited hosting capacity (Murray et al., 2021). Fast PV fluctuations and changing load conditions can cause frequent voltage deviations that conventional equipment such as tap changers and capacitor banks cannot correct due to their slower operating times (Bu et al., 2022; Lazo and Watts, 2024). As a result, maintaining voltage quality in ADNs requires faster, adaptive control strategies located closer to the grid edge (Li et al., 2021).

Inverter-based Volt/Var Control (VVC) has become the primary mechanism for regulating voltage in PV-rich feeders and is now required by modern grid codes (Abdelkader et al., 2024; Gui et al., 2023). Operating on local voltage measurements, it enables fast, continuous voltage support without centralized coordination (Hashemnezhad and Aristidou, 2025). However, reactive power capability is fundamentally constrained by the inverter's apparent power limit (Hamdan et al., 2023). When PV output is high, the reactive margin becomes limited due to the capability curve (Weckx et al., 2014), and voltage violations may persist even if VVC is fully engaged. Similarly, during high-load conditions with low PV, limited reactive support can restrict undervoltage correction (Aboshady et al., 2023). In such cases, VVC alone is insufficient, motivating the use of a supplementary control resource.

Active power flexibility provides an additional means of voltage regulation when reactive support reaches its limits. Several works have explored coordinating VVC with Demand Response (DR) to enhance voltage profiles or reduce operational costs. For example, risk-aware VVC schemes jointly schedule inverter reactive power, storage systems, and DR to manage uncertainty in PV-rich feeders (Gholami et al., 2023). Co-optimization approaches have been proposed to simultaneously schedule active and reactive power under PV uncertainty using chance-constrained or stochastic formulations (Najafi and Livani, 2025). Real-time coordination of active and reactive power has also been explored using model-free inverter control strategies (Nazaripouya et al., 2019) and distributed optimization to extend voltage regulation capabilities in high R/X feeders (Hu et al., 2020). Moreover, flexible loads have been incorporated into two-stage dispatch frameworks to complement inverter VVC and improve operational reliability (Zhang et al., 2020). More recently, hierarchical and bi-level optimization schemes have coordinated DR with optimal inverter VVC to increase renewable hosting capacity (Zenhom et al., 2025) or to tune adaptive droop characteristics (MansourLakouraj et al., 2023).

Although these studies explored the joint use of VVC and DR to improve voltage profiles, reduce losses, or manage uncertainty in PV-rich feeders, existing approaches typically rely on centralized optimization, multi-interval scheduling, or continuous activation of flexible loads. While effective, these methods suffer from two key shortcomings: flexibility may be activated even when reactive power alone would have been sufficient—resulting in unnecessary customer discomfort or increased operational cost—and they do not explicitly identify when inverter reactive capability has become saturated, which is the critical condition under which active power support is actually required.

⋆ This project has received funding from the European Union's Horizon Europe Framework Programme (HORIZON) under the GA n. 101120278 - DENSE.

This paper addresses these gaps by proposing a real-time hierarchical voltage control framework that prioritizes reactive power from distributed inverters and activates aggregated demand flexibility only when reactive headroom is insufficient. The approach integrates three key components: a Multi-Agent Reinforcement Learning (MARL) layer that learns decentralized inverter reactive power policies for primary voltage regulation; a lightweight Graph Neural Network (GNN) power-flow estimator that rapidly predicts post-control voltages and enables real-time detection of remaining violations without repeatedly solving full AC power flow; and a single aggregator agent equipped with an activation gate that adjusts active power only when voltage limits remain violated and the local inverters operate at their capability boundaries. The contributions of this paper are as follows:

- A hierarchical voltage control scheme in which inverter reactive power is used first, and aggregated demand flexibility is activated only when reactive headroom is insufficient.
- A lightweight GNN power-flow estimator enabling fast voltage prediction after each reactive power action without running full AC power flow.
- A reinforcement learning framework combining multi-agent inverter control with a single aggregator agent for coordinated active power adjustments, guided by an activation mechanism that encourages minimal use of flexibility.
- Case studies on a high-PV distribution feeder showing that the proposed method preserves voltage quality under normal conditions and successfully mitigates severe overvoltage and undervoltage with only limited (10–15%) flexible load adjustments.

The remainder of the paper is structured as follows. Section 2 describes the voltage regulation problem. The proposed learning-based frameworks are detailed in Section 3. Section 4 presents the numerical results. The paper concludes with Section 5.

2. PROBLEM FORMULATION

Ensuring voltage stability in ADNs with high PV penetration requires coordinating reactive and active power resources. Traditional devices such as tap changers and capacitor banks are too slow to handle fast, localized PV fluctuations, whereas smart inverters and flexible loads provide real-time control. In this work, voltage regulation is formulated using inverter reactive power as the primary resource and aggregated demand flexibility as a secondary input. The objective is to keep all bus voltages within the 0.95–1.05 p.u. limits while minimizing active power losses, activating flexibility only when reactive capability is insufficient.

2.1 Network Model

Consider a radial distribution network with buses $\mathcal{N} = \{1, 2, \ldots, N\}$ and lines $\mathcal{E} \subseteq \mathcal{N} \times \mathcal{N}$. The steady-state active and reactive power flows from bus i to bus j are

$$p_{ij} \;=\; g_{ij} v_i^2 \;-\; v_i v_j \big(g_{ij} \cos\delta_{ij} + b_{ij} \sin\delta_{ij}\big), \qquad (1)$$

$$q_{ij} \;=\; -\, b_{ij} v_i^2 \;+\; v_i v_j \big(g_{ij} \sin\delta_{ij} - b_{ij} \cos\delta_{ij}\big), \qquad (2)$$

where v_i and v_j are the voltage magnitudes at buses i and j, $\delta_{ij} = \delta_i - \delta_j$ is the voltage angle difference, and g_{ij}, b_{ij} denote the conductance and susceptance of the line connecting the two buses, respectively.

The total active power loss at time t is

$$P_{\text{loss}}(t) \;=\; \sum_{(i,j)\in\mathcal{E}} g_{ij} \left(v_i^2 + v_j^2 - 2 v_i v_j \cos\delta_{ij}\right). \qquad (3)$$

At each bus i, the active and reactive power balances are maintained as

$$P_i^{\text{G}}(t) - P_i^{\text{L}}(t) \;=\; \sum_{j\in\mathcal{N}_i} p_{ij}(t), \qquad (4)$$

$$Q_i^{\text{G}}(t) - Q_i^{\text{L}}(t) \;=\; \sum_{j\in\mathcal{N}_i} q_{ij}(t), \qquad (5)$$

where P_i^{G} and Q_i^{G} denote the active and reactive power injections from PV inverters, P_i^{L} and Q_i^{L} represent the active and reactive load demands, and $\mathcal{N}_i$ is the set of neighboring buses connected to node i. The network must operate within technical voltage limits:

$$V_{\min} \;\le\; v_i(t) \;\le\; V_{\max}, \quad \forall i \in \mathcal{N}, \qquad (6)$$

These equations form the foundation of the voltage regulation problem, where inverters and flexible loads are used as controllable resources to minimize voltage deviations and network losses while respecting operational constraints.

2.2 Inverter-Based Volt/Var Control

PV inverters are capable of providing reactive power support for local voltage regulation in ADNs. Given the apparent power limit $S_i^{\max}$ of inverter i, the active and reactive power outputs $p_i(t)$ and $q_i(t)$ must satisfy the capability constraint

$$p_i^2(t) + q_i^2(t) \le \left(S_i^{\max}\right)^2, \quad \forall i \in \mathcal{N}. \qquad (7)$$

Under normal operation, the inverter adjusts its reactive power output according to a control strategy aimed at minimizing voltage deviations and network power losses while respecting (7). The Volt/Var Optimization (VVO) problem can therefore be formulated as

$$\begin{aligned} \min_{\{q_i(t)\}} \;& J(t) = \sum_{(i,j)\in\mathcal{E}} g_{ij} \left(v_i^2 + v_j^2 - 2 v_i v_j \cos\delta_{ij}\right) \\ \text{s.t.} \;& V_{\min} \le v_i(t) \le V_{\max}, \quad \forall i \in \mathcal{N}, \\ & Q_i^{\min} \le q_i(t) \le Q_i^{\max}, \quad \forall i \in \mathcal{N}, \\ & \textit{Power flow equations} \end{aligned} \qquad (8)$$

where $J(t)$ represents the total active power loss at time t, and $Q_i^{\min}$ and $Q_i^{\max}$ denote the inverter reactive power limits. This formulation captures the steady-state voltage regulation problem when only inverter-based reactive power is utilized. However, when the available reactive power margin is depleted (i.e., the inverter operates near the boundary of (7)), its ability to correct voltage deviations becomes limited. In such cases, additional active power flexibility may be required to restore voltages within the permissible range.

2.3 Aggregated Demand Flexibility Model

When inverter reactive power support reaches its capability limit, additional active power flexibility from control-

lable loads can provide a complementary voltage regulation mechanism. In this work, fast real-time flexibility is assumed to originate from devices capable of rapid power modulation, such as residential HVAC systems, electric water heaters, electric vehicle fleets, and small commercial loads with short-term power adjustability. Flexible demand resources are aggregated and represented as controllable active power adjustments $\Delta P^{\mathrm{L}}(t)$ at the distribution network.

Let $P^{\mathrm{L},0}(t)$ denote the baseline active power demand and $\Delta P^{\mathrm{L}}(t)$ the controllable variation. The actual active power consumption is therefore

$$P^{\mathrm{L}}(t) = P^{\mathrm{L},0}(t) + \Delta P^{\mathrm{L}}(t), \tag{9}$$

with the flexibility constrained by the available range of the load aggregator

$$-\beta P^{\mathrm{L},0}(t) \le \Delta P^{\mathrm{L}}(t) \le \beta P^{\mathrm{L},0}(t), \tag{10}$$

where $\beta \in [0,1]$ denotes the proportion of demand that can be adjusted. Typical flexibility levels considered in this study range between 5% and 20% of the baseline load. The active power adjustment directly influences bus voltages through the network resistance-to-reactance ratio (R/X). In feeders with high R/X values, changes in active power can noticeably affect voltage magnitudes. The objective of the flexibility control layer is to modify $\Delta P^{\mathrm{L}}(t)$ only when voltage violations persist and reactive power resources are fully utilized.

To capture this conditional activation, a binary indicator variable $\phi_i(t)$ is introduced:

$$\phi_i(t) = \begin{cases} 1, \text{ if } |v_i(t)| \notin [V_{\min}, V_{\max}] \text{ and } |q_i(t)| = Q_i^{\max}, \\ 0, \text{ otherwise.} \end{cases} \tag{11}$$

Hence, demand flexibility is triggered only when both a voltage limit violation and reactive power saturation occur. This selective mechanism ensures that active power adjustments are applied sparsely, preserving customer comfort and minimizing unnecessary power variations.

2.4 Hierarchical Coordination Principle

The proposed scheme uses two layers executed at each control interval t: (i) a primary VVC layer that adjusts inverter reactive power, and (ii) a secondary flexibility layer that modifies active demand *only if* reactive headroom is exhausted and voltage limits are violated.

First, solve the primary VVO (8) to obtain $q_i^{\star}(t)$ and the resulting voltages $v_i(t)$. Then evaluate the local activation flags $\phi_i(t)$ as in (11) and define a feeder-level trigger

$$\Phi(t) = \max_{i \in \mathcal{N}} \phi_i(t) \in \{0,1\}$$

If $\Phi(t) = 1$, activate the secondary problem that computes the smallest demand adjustments needed to restore voltages within limits. Let the adjustable load at bus i be constrained by (10); the gating condition is enforced by tightening the bounds with $\phi_i(t)$:

$$-\phi_i(t)\,\beta P^{\mathrm{L},0}(t) \le \Delta P^{\mathrm{L}}(t) \le \phi_i(t)\,\beta P^{\mathrm{L},0}(t) \tag{12}$$

so that $\Delta P^{\mathrm{L}}(t) = 0$ whenever $\phi_i(t) = 0$. The flexibility dispatch is then obtained from

$$\begin{aligned} \min_{\{\Delta P^{\mathrm{L}}(t)\}} \quad & \alpha \left\| \Delta P^{\mathrm{L}}(t) \right\|_1 \\ \text{s.t.} \quad & V_{\min} \le v_i(t) \le V_{\max}, \quad \forall i \in \mathcal{N}, \\ & \textit{Bounds } (12), \\ & \textit{Power flow equations} \end{aligned} \tag{13}$$

The parameter $\alpha > 0$ penalizes flexibility usage, encouraging sparse, minimal adjustments. When $\Phi(t) = 0$, (13) admits only the trivial solution, and the secondary layer is therefore skipped.

3. LEARNING-BASED COORDINATION FRAMEWORK

To enable real-time coordination of inverter reactive power and conditional demand flexibility, we adopt a learning-based control architecture combining MARL and a fast GNN surrogate. MARL agents learn local reactive power policies, while the GNN rapidly predicts post-action voltages to identify any remaining violations. If limits are exceeded, the flexibility controller adjusts active power accordingly. This approach enables efficient, scalable voltage regulation without repeatedly solving full AC power flows.

3.1 MARL for VVC

The lower control layer is composed of inverter agents that learn reactive power control policies to maintain voltage stability. Each inverter agent i observes its local operating state $s_i(t) = [v_i(t), p_i(t), q_i(t), v_i^{\mathrm{ref}}]$ and selects a continuous action $a_i(t) = q_i^{\mathrm{ref}}(t)$ within its reactive power capability limit (7). The joint action $\mathbf{a}(t) = [a_1(t), a_2(t), \ldots, a_N(t)]$ is applied to the network, after which a fast voltage estimation is performed using the GNN surrogate model described in Subsection 3.2.

The GNN outputs the estimated voltage magnitudes $\hat{v}_i(t{+}1)$ corresponding to the selected reactive power actions. If all estimated voltages satisfy $V_{\min} \le \hat{v}_i(t{+}1) \le V_{\max}$ for all $i \in \mathcal{N}$, the control episode proceeds normally. Otherwise, the system identifies that inverter reactive capability is insufficient, and the *aggregator agent* is activated to apply an active power adjustment $\Delta P^{\mathrm{L}}(t)$ according to the flexibility model in (9)–(11).

The reward function guiding the MARL training is designed to enforce voltage regulation within acceptable limits, discourage excessive power losses, and minimize unnecessary demand flexibility activation. A piecewise voltage penalty function $f_v(\cdot)$ is introduced:

$$f_v(v_i(t)) = \begin{cases} 0, & V_{\min} \le v_i(t) \le V_{\max}, \\ \alpha_v \left(v_i(t) - V_{\max}\right)^2, & v_i(t) > V_{\max}, \\ \alpha_v \left(V_{\min} - v_i(t)\right)^2, & v_i(t) < V_{\min}, \end{cases} \tag{14}$$

which penalizes only the magnitude of voltage violations outside the acceptable range $[V_{\min}, V_{\max}]$. To ensure that demand flexibility is used sparingly, a smooth penalty function $f_{\mathrm{flex}}(t)$ is employed to penalize the magnitude of load adjustments rather than merely their activation:

$$f_{\mathrm{flex}}(t) = \alpha_f \left\| \Delta P^{\mathrm{L}}(t) \right\|_1, \tag{15}$$

Algorithm 1 MARL with Demand Flexibility

```
Initialize actor and critic networks for inverter agents
Initialize replay buffer B
for each training episode do
  for each time step t do
    Observe local states s_i(t) for all inverter agents
    Select actions a_i(t) = π_i(s_i(t)) + N_t
    Apply reactive power actions to the network
    Estimate voltages v̂_i(t+1) using the GNN surrogate
    if any v̂_i(t+1) ∉ [V_min, V_max] then
      Trigger flexibility adjustment ΔP^L(t)
    end if
    Compute reward r(t) using (16)
    Store (s, a, r, s') in B
    Update actor and critic networks from sampled minibatch
  end for
end for
```

where $\Delta P^{\mathrm{L}}(t)$ is the vector of load adjustments at time t, and α_f weights the cost of modifying consumption. The overall reward at time step t is then given by

$$r(t) = -\left[\sum_{i\in\mathcal{N}} f_v\big(v_i(t{+}1)\big) + P_{\mathrm{loss}}(t) + f_{\mathrm{flex}}(t)\right], \quad (16)$$

This formulation encourages the agent to (i) keep voltages within limits, (ii) reduce power losses, and (iii) employ the minimum amount of demand flexibility needed for restoring voltage security. All inverter agents are trained using a centralized training and decentralized execution framework based on the Multi-Agent Twin Delayed Deep Deterministic Policy Gradient (MATD3) algorithm.

3.2 GNN for Fast Power Flow Estimation

Conventional reinforcement learning frameworks require frequent evaluation of the power flow equations to assess the impact of control actions on network voltages. However, iterative solvers such as Newton–Raphson or backward–forward sweep are computationally expensive and may slow down both training and real-time operation. To overcome this limitation, a GNN surrogate model is employed to approximate the nonlinear power flow mapping with high accuracy and low computational cost.

The distribution network is represented as a directed graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$, where each node corresponds to a bus and each edge represents a line or transformer. Each node i is characterized by features $x_i = [P_i^{\mathrm{G}}, Q_i^{\mathrm{G}}, P_i^{\mathrm{L}}, Q_i^{\mathrm{L}}, v_i]$, while edge features e_{ij} include the line parameters (r_{ij}, x_{ij}). The objective of the GNN is to learn a mapping $f_\theta : (x_i, e_{ij}) \to \hat{v}_i$, parameterized by weights θ, that predicts the nodal voltage magnitudes $\hat{v}_i$.

At each message-passing iteration k, node embeddings are updated as

$$h_i^{(k+1)} = \sigma\left(W_1 h_i^{(k)} + \sum_{j\in\mathcal{N}_i} \phi\Big(h_i^{(k)}, h_j^{(k)}, e_{ij}\Big)\right), \quad (17)$$

where $h_i^{(k)}$ denotes the hidden representation of node i at layer k, $\phi(\cdot)$ is a learnable message function, W_1 is a trainable weight matrix, and $\sigma(\cdot)$ is a nonlinear activation function (e.g., ReLU). After K iterations, the predicted voltage magnitude at node i is obtained as

$$\hat{v}_i = W_2 h_i^{(K)} + b, \quad (18)$$

where W_2 and b are the final layer parameters. The network is trained offline using supervised learning, with ground-truth voltages obtained from power flow simulations over diverse operating conditions and inverter control actions. The loss function is defined as

$$\mathcal{L} = \frac{1}{N}\sum_{i=1}^{N}\left(v_i^{\mathrm{PF}} - \hat{v}_i\right)^2, \quad (19)$$

where v_i^{PF} is the true voltage magnitude from the numerical power flow and $\hat{v}_i$ is the GNN prediction.

According to Algorithm 1, the GNN is embedded into the MARL loop to provide fast voltage estimation after each reactive power action. Specifically, following inverter control actions $a_i(t) = q_i^{\mathrm{ref}}(t)$, the GNN predicts the corresponding nodal voltages $\hat{v}_i(t{+}1)$ using (17) and (18). If any of the predicted voltages violate the operational limits $[V_{\min}, V_{\max}]$, the upper-layer flexibility controller is triggered to adjust the active power demand $\Delta P^{\mathrm{L}}(t)$. This hybrid integration of learning-based voltage estimation and multi-agent control enables near real-time decision making without repeated iterative power flow calculations.

4. NUMERICAL ANALYSIS

This section evaluates the proposed control framework through numerical experiments that assess its ability to coordinate inverter-based reactive power control with conditional demand flexibility activation. The analysis focuses on three aspects: (i) the behavior of the VVC under normal operating conditions, (ii) the effectiveness of flexibility when reactive power becomes insufficient, and (iii) the combined impact of both resources on voltage regulation and power losses. The following subsections outline the dataset and training setup, followed by a detailed discussion of the resulting voltage profiles, flexibility utilization, and comparative performance across operating scenarios.

4.1 Dataset and Training Setup

This study evaluates the proposed hierarchical control scheme using a modified IEEE 33-bus distribution network as shown in Figure 1. The load profiles are obtained from three years of Portuguese electricity consumption data (Trindade, 2015), and PV generation profiles are sourced from the Elia Group real-time solar dataset reported in (Elia Group, 2024). Both time series are interpolated to a 3-minute resolution in order to capture rapid fluctuations in PV output and demand and to facilitate high-resolution control actions during training. These datasets are mapped onto the IEEE 33-bus system by assigning load and PV patterns to selected buses proportionally to their nominal loading levels.

For training the MARL controller, each episode spans an entire 24-hour operational cycle comprising 480 time steps. The GNN surrogate model is trained offline on a dataset constructed from 15,000 power flow simulations generated by randomly sampling load multipliers, PV scaling factors, and inverter reactive power actions. The resulting voltage magnitudes serve as ground truth labels for learning.

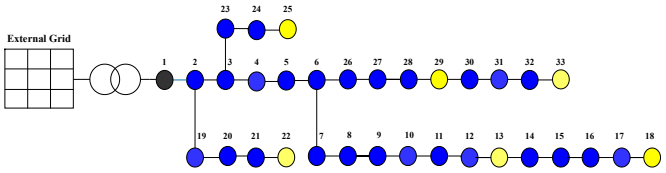


Fig. 1. Modified IEEE 33-bus distribution network

Table 1. MARL Training Hyperparameters

Parameter	Value
Learning Rate	0.001
Discount Factor (γ)	0.99
Batch Size	32
Replay Buffer Size	5×10^3
Target Network	$\tau = 0.1$
Gradient Clipping	$\epsilon = 1.0$
Training Episodes	1000

Table 2. GNN Training Hyperparameters

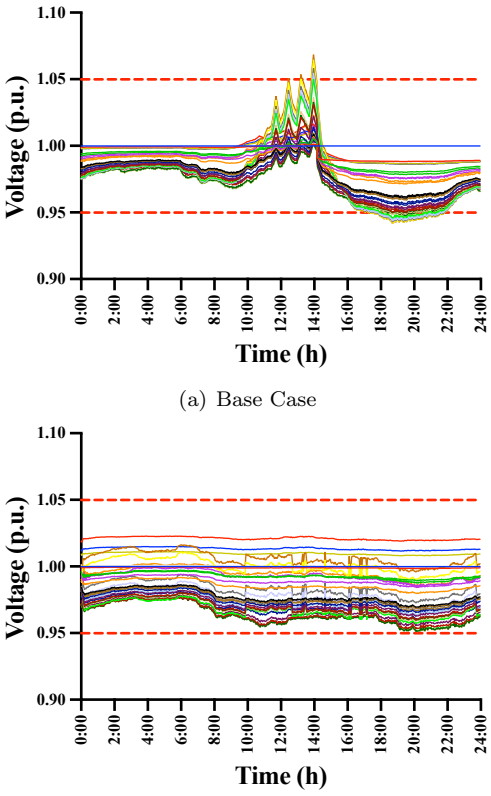
Parameter	Value
Graph Layers	3 Graph Convolution Layers
Hidden Units per Layer	64
Activation Function	ReLU
Learning Rate	0.001
Batch Size	64
Epochs	100

Hyperparameters for both the MARL framework and the GNN model are summarized in Tables 1 and 2. They were selected through empirical tuning, prior studies, and stability considerations to ensure efficient learning, fast convergence, and robust performance across varying conditions. All experiments were run on a MacBook Pro with an Apple M4 Max chip, 36-core GPU, and 36 GB unified memory, using macOS Sonoma. Each training episode simulates a full day at 3-minute intervals (480 steps). Training for 1,000 episodes took 49 minutes.

4.2 Results

In this subsection, the proposed hierarchical voltage control scheme is evaluated through case studies that examine its behavior under different operating conditions and flexibility levels. The first experiment considers a mild scenario in which PV generation and load variations induce only moderate voltage deviations. As shown in Fig. 2, the MARL-based VVC maintains all bus voltages within the permissible range without activating the demand flexibility layer. In this case, the inverter reactive power capability is sufficient to correct all deviations, and the activation gate remains inactive. These results confirm that the framework naturally prioritizes reactive power and refrains from using flexibility when it is not required.

Figure 3 presents the worst-case operating conditions on a sunny day leading to the overvoltage event. As shown in Fig. 3(a), PV generation peaks sharply around noon, creating a large net power injection into the feeder. This results in the significant voltage rise illustrated in Fig. 3(b), where multiple buses exceed the upper limit of 1.05 p.u. in the absence of any control action. At these hours of maximum PV generation, the inverters operate close to their rated apparent power, leaving only a small remaining margin for reactive power support.



(a) Base Case

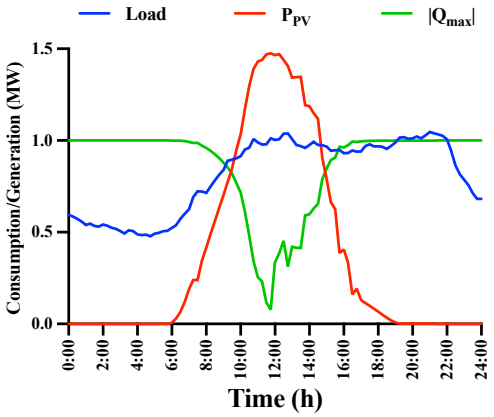
(b) Our Approach

Fig. 2. Voltage profile of all buses on a random day

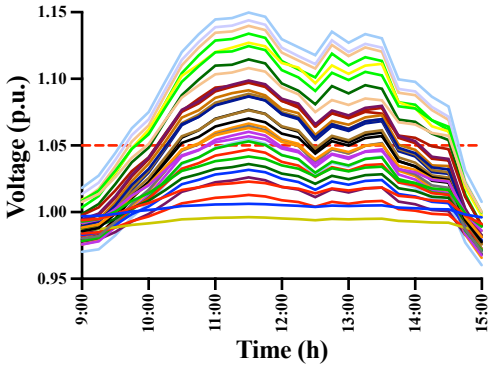
Because inverter operation must satisfy the capability constraint (7), the large active power output severely restricts the available q_i, sharply reducing the ability of the inverters to counteract the voltage rise.

Figure 4 shows the reactive power control layer reaction to the overvoltage by driving each inverter to inject reactive power. However, as shown in Fig. 4(a) the required reactive support $|Q_{\text{need}}|$ during the noon hours exceeds the instantaneous inverter capability $|Q_{\max}|$, which is severely reduced due to high active power generation. Consequently, although the VVC mechanism partially mitigates the voltage rise, Fig. 4(b) shows that several buses remain above the 1.05 p.u. limit. This confirms that reactive power alone is insufficient under this stressed operating condition, necessitating the activation of the demand flexibility layer to fully restore voltage compliance.

When the remaining overvoltage could not be eliminated through VVC alone, the demand-flexibility layer was activated. Table 3 summarizes the resulting performance under different flexibility limits. With only 5% available flexibility, the controller lacks sufficient corrective capability, resulting in a residual deviation of 1.12% and a peak voltage of 1.065 p.u. Increasing the limit to 10% significantly improves performance, while 15% flexibility is fully effective, restoring all voltages within the permissible band. Notably, the maximum flexibility actually used saturates at about 14.3% even when more (20%) is offered, owing to the reward structure, which penalizes unnecessary load adjustments. Thus, the agent applies only the minimal modification required for voltage restoration. The corresponding voltage profiles in Fig. 5 confirm that the



(a) Load, PV generation, and available reactive power



(b) Voltage profiles in the base case (no control).

Fig. 3. Load–PV conditions and resulting voltages in the overvoltage scenario.

Table 3. Performance under different flexibility limits

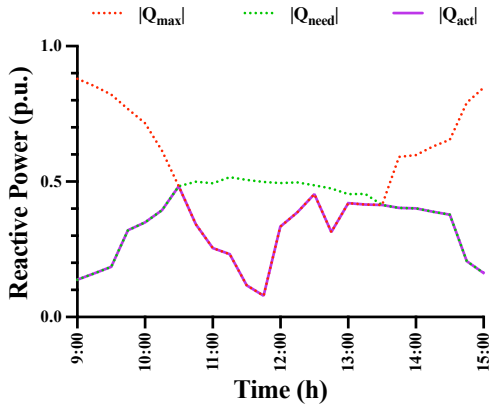
Flexibility limit (%)	5%	10%	15%	20%
Max Flexibility Used (%)	5	10	14.33	14.34
Voltage Deviation (%)	1.12	0.46	0	0
Max Voltage (p.u.)	1.065	1.053	1.049	1.049

Table 4. Performance of the proposed approach

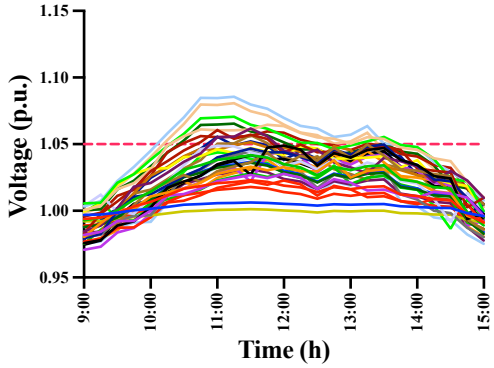
	Base case	MARL	MARL-FLEX
Max Flexibility Used (%)	–	–	9.67
Voltage Deviation (%)	4.2	0.78	0
Min voltage (p.u.)	0.932	0.943	0.952
Max voltage (p.u.)	1.028	1.017	1.017
Power loss (%)	6.13%	3.92%	3.38%

coordinated VVC and flexibility control successfully bring all bus voltages below 1.05 p.u.

Finally, the proposed approach is evaluated under a severe undervoltage scenario, with results summarized in Table 4. In the base case, the feeder experiences significant voltage drops, reaching a minimum of 0.932 p.u. Pure MARL-based VVC mitigates the issue by injecting reactive power where available, improving the minimum voltage to 0.943 p.u., but the high-load and low-PV conditions leave insufficient reactive capability to fully restore voltage levels. When demand flexibility is incorporated (MARL–FLEX), the controller selectively reduces loads at key buses, raising the minimum voltage to 0.952 p.u. and eliminating all violations while using only 9.67% of



(a) Reactive power capability, requirement, and actual injection during the overvoltage.



(b) Voltages after reactive power control (no flexibility yet).

Fig. 4. Voltage profile after Volt/Var control under the overvoltage condition.

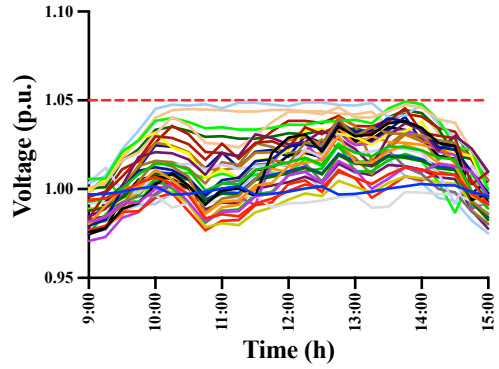


Fig. 5. Voltage profile with coordinated demand flexibility activation during the overvoltage.

the available flexibility. This minimal intervention aligns with the reward structure, which penalizes unnecessary consumption changes. Moreover, MARL–FLEX achieves lower power losses (3.38%), since reducing demand decreases current flows and corresponding losses (3). These results demonstrate that the hierarchical framework effectively coordinates reactive power and demand flexibility to maintain voltage stability under stressed operating conditions.

5. CONCLUSION

This paper presented a hierarchical voltage control framework that combines MARL-based VVC with conditionally activated demand flexibility. Using a GNN-based power flow estimator and a trigger mechanism, the controller activates flexibility only when inverter reactive power is fully utilized. Tests on a modified IEEE 33-bus feeder with realistic PV and load profiles showed that, under normal conditions, VVC alone maintains acceptable voltages without relying on flexibility. In severe overvoltage and undervoltage scenarios, where reactive power capability becomes insufficient, the controller activates only the minimal amount of flexibility needed to restore voltage compliance. The results confirm that the proposed approach provides robust, efficient voltage regulation while avoiding unnecessary changes to customer consumption.

ACKNOWLEDGEMENTS

This project has received funding from the European Union's Horizon Europe Framework Programme (HORIZON) under the GA n. 101120278 - DENSE.

REFERENCES

Abdelkader, S.M., Kinga, S., Ebinyu, E., Amissah, J., Mugerwa, G., Taha, I.B., and Mansour, D.E.A. (2024). Advancements in data-driven voltage control in active distribution networks: A comprehensive review. *Results in Engineering*, 23, 102741.

Aboshady, F., Pisica, I., Zobaa, A.F., Taylor, G.A., Ceylan, O., and Ozdemir, A. (2023). Reactive power control of pv inverters in active distribution grids with high pv penetration. *IEEE Access*, 11, 81477–81496.

Bu, S., Meegahapola, L.G., Wadduwage, D.P., and Foley, A.M. (2022). Stability and dynamics of active distribution networks (adns) with d-pmu technology: A review. *IEEE Transactions on Power Systems*, 38(3), 2791–2804.

Elia Group (2024). Solar pv power generation data [dataset]. Elia Open Data Portal. URL `https://www.elia.be/en/grid-data/power-generation/solar-pv-power-generation-data`.

Gholami, K., Azizivahed, A., Arefi, A., and Li, L. (2023). Risk-averse volt-var management scheme to coordinate distributed energy resources with demand response program. *International Journal of Electrical Power & Energy Systems*, 146, 108761.

Gui, Y., Nainar, K., Bendtsen, J.D., Diewald, N., Iov, F., Yang, Y., Blaabjerg, F., Xue, Y., Liu, J., Hong, T., et al. (2023). Voltage support with pv inverters in low-voltage distribution networks: An overview. *IEEE Journal of Emerging and Selected Topics in Power Electronics*, 12(2), 1503–1522.

Hamdan, I., Alfouly, A., and Ismeil, M.A. (2023). A literature review on hosting capacity methodologies and inverter control technologies for photovoltaic system. In *2023 IEEE Conference on Power Electronics and Renewable Energy (CPERE)*, 1–7. IEEE.

Hashemnezhad, M. and Aristidou, P. (2025). Safe multi-agent reinforcement learning volt-var optimization in active distribution networks. In *Proc. IEEE PES Innovative Smart Grid Technologies Europe (ISGT Europe)*. Valletta, Malta.

Hu, X., Liu, Z.W., Wen, G., Yu, X., and Liu, C. (2020). Voltage control for distribution networks via coordinated regulation of active and reactive power of dgs. *IEEE Transactions on Smart Grid*, 11(5), 4017–4031.

Lazo, J. and Watts, D. (2024). Stochastic model for active distribution networks planning: An analysis of the combination of active network management schemes. *Renewable and Sustainable Energy Reviews*, 191, 114156.

Li, X., Li, H., Li, S., Jiang, Z., and Ma, X. (2021). Review on reactive power and voltage optimization of active distribution network with renewable distributed generation and time-varying loads. *Mathematical Problems in Engineering*, 2021(1), 1196369.

MansourLakouraj, M., Hosseinpour, H., Livani, H., Benidris, M., and Fadali, M.S. (2023). Volt/var support and demand response co-optimization in distribution systems with adaptive droop control of inverters. In *2023 IEEE Texas Power and Energy Conference (TPEC)*, 1–6. IEEE.

Murray, W., Adonis, M., and Raji, A. (2021). Voltage control in future electrical distribution networks. *Renewable and Sustainable Energy Reviews*, 146, 111100.

Najafi, S. and Livani, H. (2025). Chance-constrained co-optimization of demand response and volt/var under gaussian mixture model uncertainty. *Renewable Energy Focus*, 53, 100674.

Nazaripouya, H., Pota, H.R., Chu, C.C., and Gadh, R. (2019). Real-time model-free coordination of active and reactive powers of distributed energy resources to improve voltage regulation in distribution systems. *IEEE Transactions on Sustainable Energy*, 11(3), 1483–1494.

Trindade, A. (2015). Electricityloaddiagrams20112014 [dataset]. UCI Machine Learning Repository. doi:10.24432/C58C86. URL `https://archive.ics.uci.edu/ml/datasets/ElectricityLoadDiagrams20112014`.

Weckx, S., Gonzalez, C., and Driesen, J. (2014). Combined central and local active and reactive power control of pv inverters. *IEEE Transactions on Sustainable Energy*, 5(3), 776–784.

Zenhom, Z.M., Aleem, S.H.A., El Zahab, E.E.D.A., Ali, Z.M., Almalaq, A., and Boghdady, T.A. (2025). Synergistic hierarchical bi-level optimization of demand response programs and smart inverter volt/var control for maximizing hosting capacity of power systems with renewable-based distributed generators. *IEEE Access*.

Zhang, Y., Song, X., Li, Y., Zeng, Z., Yong, C., Sidorov, D., and Lv, X. (2020). Two-stage active and reactive power coordinated optimal dispatch for active distribution network considering load flexibility. *Energies*, 13(22), 5922.