RESEARCH ARTICLE OPEN ACCESS

Spatio-Temporal Graph Transformer MARL for Resilient Point of Common Coupling Control in Networked Grid-Tied Micro-grids

Nicholas Nyaika[1*], Cosmas Mwikirize (PhD)[2], Andrew Katumba (PhD)[3]

[1]Department of Electrical and Computer Engineering/ Makerere University, Kampala, Uganda, | ORCID ID; 0009-0007-5901-6102

[2]Department of Electrical and Computer Engineering/ Makerere University, Kampala, Uganda, | ORCID ID; 0000-0002-3393-2758

[3] Department of Electrical and Computer Engineering/ Makerere University, Kampala, Uganda, | ORCID ID; 0000-0001-7699-8426

***Corresponding author email:** nyaikanick@gmail.com*

Abstract — The increasing penetration of renewable energy sources introduces significant variability and uncertainty into grid-connected microgrids (MGs), affecting voltage stability, frequency regulation, power quality, fault ride-through capability, and operational resilience at the Point of Common Coupling (PCC). These challenges become more pronounced in networked MGs with changing electrical and communication topologies, heterogeneous operating conditions, communication constraints, and variable loads. This study proposes an adaptive, distributed control framework that integrates a Spatio-Temporal Graph Transformer (ST-GT) for learning evolving network relationships, cooperative Multi-Agent Reinforcement Learning (MARL) for decentralized control with physics-informed constraints, and differential-privacy-enhanced federated selective sharing for secure collaboration among MGs. The framework is evaluated on modified IEEE 33-bus and IEEE 69-bus networked MG systems using high-fidelity digital twins implemented in pandapower and pymgrid. Experiments consider renewable penetration of 50% to 80%, short-circuit ratios below 3, communication failures, forecast uncertainty, islanding, and fault ride-through disturbances. Compared with benchmark control approaches, the proposed framework reduces PCC voltage deviations and total harmonic distortion by 26% to 33%, achieves fault ride-through success rates above 96%, improves economic performance by 17% to 23%, and maintains robust operation under forecast errors exceeding 40%. These findings demonstrate that the framework provides a scalable, resilient, and privacy-preserving approach for coordinated control of networked grid-connected MGs.

Keywords — Point of Common Coupling, Networked MGs, Spatio-Temporal Graph Transformer, Multi-Agent Reinforcement Learning.

I. INTRODUCTION

The rapid transition to sustainable and low-carbon energy systems has accelerated the integration of Renewable Energy Sources (RES), Battery Energy Storage Systems (BSS), Electric Vehicles (EVs), and controllable loads into networked MGs. While these systems improve energy access, resilience, and renewable utilization, they also introduce significant operational challenges at the PCC where local energy resources interact with the utility grid. Increasing renewable penetration creates substantial uncertainty that is made worse by load fluctuations, EV charging behaviour, and market dynamics. Also, this has made voltage regulation, frequency stability, harmonic mitigation, fault response, and bidirectional power flow management increasingly complex [5].

Conventional PCC control strategies based on hierarchical droop control, secondary frequency restoration, and tertiary economic dispatch are limited by their reliance on simplified system models and predefined operating assumptions. Their performance deteriorates under renewable uncertainty, communication delays, reduced system inertia, and dynamically changing network topologies [8]. Although MPC improves constraint handling and predictive optimization, its computational complexity grows rapidly with system size and uncertainty, limiting scalability for large networked MGs. RL, MARL, and Deep Reinforcement Learning (DRL) have emerged as promising alternatives because they enable adaptive control through interaction with the environment. However, existing MARL approaches often lack physical interpretability, struggle with dynamic network topologies, and do not adequately address privacy concerns in collaborative learning environments.

Recent advances in GNNs and Graph Transformers (GTs) provide powerful tools for representing power system topologies and capturing spatial dependencies among interconnected assets. In particular, GTs can model both spatial and temporal relationships while adapting to evolving electrical and communication networks [4, 1]. Nevertheless, existing graph-based control approaches rarely incorporate power system physics, long-range temporal dependencies, privacy preservation, and decentralized coordination within a unified framework. Furthermore, traditional reward-based RL cannot guaranty compliance with critical operational constraints such as power flow equations, voltage limits, frequency stability requirements, and harmonic distortion thresholds [3].

To address these challenges, this paper proposes a novel framework that combines dynamic graph learning, physics informed control, decentralized multi-agent coordination, communication-efficient FL, and privacy-preserving mechanisms within a unified architecture. Consequently, it enables resilient, scalable, and privacy-aware operation under renewable uncertainty, communication failures, weak-grid conditions, and dynamically changing network topologies.

The major contributions of this work are summarized as follows:

1) A novel ST-GT architecture is developed to model dynamically evolving electrical and communication network topologies while capturing long-range spatio-temporal dependencies arising from renewable generation variability, load uncertainty, and network reconfiguration events.

2) A Physics-Informed Multi-Agent Reinforcement Learning (PI-MARL) framework is proposed in which AC power flow equations, voltage stability constraints, swing dynamics, harmonic distortion limits, and Lyapunov-based stability conditions are directly embedded into the learning objective, thereby enabling safe and physically consistent control policies.

3) A Federated Selective Sharing (FSS) mechanism is introduced that exchanges only high-relevance graph attention parameters among participating MGs, significantly reducing communication overhead while preserving local model personalization and decentralized decision-making capabilities.

4) An adaptive Differential Privacy (DP) framework is developed to provide formal privacy guaranties through relevance-aware noise injection into selectively shared parameters, thereby mitigating information leakage while maintaining control performance.

5) A comprehensive PCC-centric optimization framework is established that simultaneously addresses voltage regulation, power quality enhancement, fault ride-through capability, seamless islanding and reconnection, resilience improvement, and economic operation within a unified control architecture.

6) Extensive validation is conducted using modified IEEE 33-bus and IEEE 69-bus benchmark systems together with high-fidelity digital twins under realistic operating conditions, including renewable uncertainty, communication failures, weak-grid conditions, and fault scenarios. The results demonstrate substantial improvements in stability, resilience, privacy preservation, communication efficiency, and economic performance compared with state-of-the-art control approaches.

II. RELATED WORK

RL has emerged as a promising approach for adaptive MG control due to its ability to learn optimal policies directly from system interactions without requiring explicit mathematical models. DRL algorithms such as Deep Q-Networks (DQN), Deep Deterministic Policy Gradient (DDPG), Twin delayed DDPG (TD3), Soft Actor-Critic (SAC), and Proximal Policy Optimisation (PPO) have demonstrated strong performance in voltage regulation, frequency control, battery scheduling, and energy management. Among these, PPO has gained significant attention because of its stability and suitability for continuous control problems. MARL further extends these capabilities by enabling decentralized coordination among multiple distributed energy resources. However, existing RL and MARL approaches often suffer from poor interpretability, unsafe exploration, limited scalability, and inadequate representation of network topology dynamics. Furthermore, most methods rely heavily on reward shaping and rarely incorporate physical system constraints directly into the learning process.

The use of RL-MPC algorithm in distributed energy systems combines the high-precision local optimization of MPC with the global optimization capability of RL [6]. This faces a limitation of amplified real-time computational burden, severe vulnerability to initial prediction errors, and the complex challenge of guaranteeing formal safety proofs when combining a neural network with rigid physical constraints.

Since power systems naturally exhibit graph structures, GNNs have been extensively applied to state estimation, fault detection, topology identification, forecasting, and voltage control. The integration of GNNs with RL has improved coordination and situational awareness in networked energy systems. Nevertheless, conventional GNNs are primarily designed for static graph structures and often fail to capture long range dependencies and topology changes associated with switching operations, faults, islanding events, and plug-and-play DER integration. GTs address these limitations through attention mechanisms capable of modelling both local and global interactions. More recently, ST-GT have demonstrated superior performance in forecasting and monitoring applications by jointly learning spatial and temporal dependencies. However, their application to adaptive PCC control remains largely unexplored.

A major limitation of data-driven control methods is their inability to guarantee compliance with fundamental physical laws. Physics-Informed Learning addresses this challenge by embedding governing equations and operational constraints directly into the learning process [9, 2]. Applications within power systems have demonstrated improved generalization, reliability, and robustness for forecasting, state estimation, and optimization tasks [10]. Although recent studies have incorporated physics-guided RL concepts, most existing approaches remain centralized, neglect harmonic distortion and fault ride-through

Open Access and Peer Reviewed Journal ISSN 2395-1303 https://www.ijetjournal.org/

topologies. Consequently, there remains a need for comprehensive physics-informed control frameworks specifically designed for networked MG PCC regulation.

The increasing decentralization of power systems has elevated the importance of data privacy and secure collaboration among independently operated energy assets. FL enables collaborative model training without sharing raw operational data and has been successfully applied to forecasting, energy management, and demand response applications [7]. However, conventional FL introduces significant communication overhead and remains vulnerable to information leakage through model inversion and membership inference attacks. Differential Privacy (DP) provides formal privacy guarantees through calibrated noise injection and has recently been integrated with FL in several smart-grid applications. Nevertheless, existing approaches rarely consider GT-based architectures, MARL, or adaptive privacy mechanisms that account for the importance of parameters. These limitations motivate the development of communication efficient and privacy preserving learning frameworks for future networked MGs.

The literature reveals a gap in the development of control architectures capable of simultaneously addressing dynamic network topology adaptation, renewable uncertainty, decentralized coordination, physical constraint satisfaction, communication efficiency, and privacy preservation. Existing studies typically focus on individual aspects such as MARL, GNNs, Physics-Informed Learning, FL, or Differential Privacy in isolation. To the best of the authors' knowledge, no previous work has comprehensively integrated ST-GT, PI-MARL, FSS, and adaptive DP within a single PCC control framework for networked grid-tied MGs. This gap forms the primary motivation for the proposed framework.

III. SYSTEM MODEL AND MATHEMATICAL FORMULATION

A. Networked MG Architecture

Consider a networked grid-tied MG cluster comprising N interconnected MGs connected to the utility grid through one or more PCCs. Each MG network includes; BESSs, controllable loads, EV charging stations, and inverter-based DERs.

The interconnected system is modelled as a dynamic graph $G_t = (V, E_t)$ where V and E_t denote the node set and time varying edge set, respectively. The node set is as per equation 1.

$$V = \{V_{PV}, V_{WT}, V_{BESS}, V_{LOAD}, V_{EV}, V_{PCC}\} \tag{1}$$

while edges represent electrical and communication inter-connections.

The graph topology is described by the dynamic adjacency matrix in equation 2

$$A_t = [a_{ij}(t)] \text{ with } a_{ij}(t) = f(Y_{ij}, Q_{comm}) \tag{2}$$

where Y_{ij} is the line admittance and Q_{comm} denotes communication quality.

B. State Space Representation

The global state at time step t is as per equation 3

$$s_t = \{\mathbf{V}_t, \mathbf{f}_t, \mathbf{P}_t, \mathbf{Q}_t, \mathbf{SOC}_t, \mathbf{THD}_t, \mathbf{A}_t\} \tag{3}$$

where $\mathbf{V}_t = [V_1, V_2,...,V_n]$, $\mathbf{f}_t = [f_1, f_2,..,f_n]$, $\mathbf{P}_t = [P_1, P_2,..,P_n]$, $\mathbf{Q}_t = [Q_1, Q_2,..,Q_n]$, and $\mathbf{SOC}_t = [SOC_1, SOC_2,..,SOCm]$ represent bus voltages, system frequencies, active power injections, reactive power injections, and battery state-of-charge levels, respectively.

Each agent observes a local subset of the global state $o^t_i = \phi_i(s_t)$, where $\phi_i(\cdot)$ denotes the observation mapping.

C. Action Space

The control action of agent i is as per equation 4

$$a_i^t = \{\Delta P_i, \Delta Q_i, P_{BESS,i}, M_i\} \tag{4}$$

where ΔP_i and ΔQ_i denote active and reactive power adjustments, $P_{BESS,i}$ is the battery charging/discharging power, and M_i is the operating mode satisfying M_i ∈{Grid-Connected, Islanded, Reconnection}.

D. AC Power Flow Constraints

The active and reactive power injections at bus i are obtained by equations 5 and 6.

$$P_i = V_i \sum_{j=1}^{N} V_j \, (G_{ij} cos\, \theta_{ij} + B_{ij} sin\, \theta_{ij}) \tag{5}$$

and

$$Q_i = V_i \sum_{j=1}^{N} V_j \, (G_{ij} sin\, \theta_{ij} - B_{ij} cos\, \theta_{ij}) \tag{6}$$

Where

$$\theta_{ij} = \theta_i - \theta_j \tag{7}$$

The PCC active and reactive power balance constraints are shown in equations 8 and 9

$$P_{\mathrm{PV}} + P_{\mathrm{WT}} + P_{\mathrm{BESS}} + P_{\mathrm{GRID}} = P_{\mathrm{LOAD}} + P_{\mathrm{LOSS}} \tag{8}$$

$$Q_{PV} + Q_{WT} + Q_{BESS} + Q_{GRID} = Q_{LOAD} + Q_{LOSS} \tag{9}$$

E. Operational Constraints

Voltage limits are $V_i^{min} \le V_i \le V_i^{max}$ with $0.95 \le V_i \le 1.05$ p.u. Frequency limits are $f^{min} \le f \le f^{max}$ with $49.5 \le f \le 50.5$ Hz. Battery operating limits are $SOC^{min} \le SOC_i \le SOC^{max}$.

F. Harmonic Distortion Constraints

Total Harmonic Distortion (THD) is defined is obtained by equation 10.

$$THD = \frac{\sqrt{\sum_{h=2}^{H} I_h^2}}{I_1} \times 100 \tag{10}$$

where I_h denotes the harmonic current component of order h. IEEE standards impose $THD < 5\%$, with the associated harmonic penalty in equation 11

$$L_{\mathrm{THD}} = \sum_h I_h^2 \tag{11}$$

IV. SPATIO-TEMPORAL GRAPH TRANSFORMER ENCODER

A. Graph Representation Learning

The proposed ST-GT generates graph embeddings shown in equation 12.

$$\mathbf{E}_t = STGT\,(\mathbf{X}_t, \mathbf{A}_t, \mathbf{T}_t) \tag{12}$$

Paper ID: IJET-V12I5P23

where H, A, and T denote node features, graph topology, and temporal information, respectively. Node features are represented by equation 13.

$$x_i = [V_i, f_i, P_i, Q_i, SOC_i, THD_i]. \quad (13)$$

B. Multi-Head Graph Attention

Attention coefficients are computed as shown in equation 14

$$e_{ij} = \frac{Q_i K_j^T}{\sqrt{d_k}} \quad (14)$$

With

$$Q_i = W_Q x_i, K_i = W_K x_i, V_i = W_V x_i.$$

The normalised attention weight is obtained by equation 16

$$\alpha_{ij} = \frac{\exp(e_{ij})}{\sum_k \exp(e_{ik})}, \quad (16)$$

yielding the node embedding in equation 17

$$h_i = \sum_j \alpha_{ij} V_j. \quad (17)$$

Multi-head attention is expressed as shown in equation 18

$$H_i = Concat\,(h^1{}_i, h^2{}_i \ldots, h^m{}_i). \quad (18)$$

C. Temporal Encoding

Long-range temporal dependencies are captured using Rotary Positional Embeddings (RoPE) in equation 19

$$RoPE(x_t) = R(t)x_t \quad (19)$$

where $R(t)$ denotes the temporal rotation operator.

D. Probabilistic Uncertainty Modelling

Renewable uncertainty is modelled using a Gaussian Mixture Model (GMM) shown in equation 20

$$p(x) = \sum_{k=1}^{K} \pi_k \, \mathrm{N}(x \mid \mu_k, \Sigma_k) \quad (20)$$

Subject to

$$\sum_{k=1}^{K} \pi_k = 1.$$

V. PHYSICS-INFORMED MULTI-AGENT REINFORCEMENT LEARNING

A. Partially Observable Markov Game

The PCC control problem is formulated as a cooperative Partially Observable Markov Game (POMG) as shown in equation 22

$$M = (N, S, A, O, P, R) \quad (22)$$

where N, S, A, O, P, and R denote the agent set, state space, function, respectively. The optimization objective represented by equation 23.

$$J(\pi) = \mathbb{E}\left[\sum_{t=0}^{T} \gamma^t r_t\right]. \quad (23)$$

B. Reward Function

The multi-objective reward function is shown in equations 24, 25, 26, 27 and 28.

$$r_t = w_1 r_{stab} + w_2 r_{qual} + w_3 r_{econ} + w_4 r_{res}, \quad (24)$$

where

$$r_{stab} = -\,(\,\| V_{PCC} - V_{ref} \| \,+\, \mid \Delta f \mid\,), \quad (25)$$

$$r_{qual} = -THD, \quad (26)$$

$$r_{econ} = -Cost, \quad (27)$$

and

$$r_{res} = -\,(FRT_{fail} + T_{recover}) \quad (28)$$

C. Physics-Informed Learning Objective

The physics-informed loss is defined as shown in equations 29, 30, 31, 32 and 33

$$L_{PIC} = \lambda_1 L_{PF} + \lambda_2 L_V + \lambda_3 L_{Swing} + \lambda_4 L_{THD} + \lambda_5 L_{Lyap} \quad (29)$$

where the power-flow residual is

$$L_{PF} = \| P_{inj} - B\theta \|^2 \quad (30)$$

the voltage regulation loss is

$$L_V = \sum_i |V_i - V_{\mathrm{ref}}|, \quad (31)$$

the swing-dynamics residual is

$$L_{Swing} = M\frac{d\omega}{dt} + D\omega - (P_m - P_e), \quad (32)$$

and the Lyapunov stability penalty is

$$L_{Lyap} = \max\,(0, \dot{V}). \quad (33)$$

The overall learning objective is given by equation 34

$$L_{\mathrm{Total}} = L_{\mathrm{PPO}} + \beta L_{\mathrm{PIC}} \quad (34)$$

VI. FEDERATED SELECTIVE SHARING WITH DIFFERENTIAL PRIVACY

A. Motivation

Future networked MGs are expected to consist of independently owned and operated energy systems. These entities may be unwilling to share operational data due to privacy concerns, commercial sensitivity, cybersecurity risks, and regulatory constraints.

Although Federated Learning (FL) enables collaborative model training without sharing raw data, conventional FL introduces significant communication overhead and remains vulnerable to model inversion attacks.

To address these challenges, a Differential Privacy-enhanced

B. Federated Learning Formulation

Consider a network containing M MGs. Each MG maintains local parameters θ_m updated according to equation 35

$$\theta_m^{k+1} = \theta_m^k - \eta \nabla L_m(\theta_m^k) \tag{35}$$

where L_m represents the local learning objective. Conventional federated averaging computes in equations 36 and 37.

$$\theta_G = \sum_{m=1}^{M} \frac{n_m}{N} \theta_m \tag{36}$$

Where

$$N = \sum_{m=1}^{M} n_m \tag{37}$$

Although effective, complete parameter sharing results in large communication costs and privacy vulnerabilities.

C. Attention-Based Selective Sharing

Instead of sharing all model parameters, only highly informative Graph Transformer attention parameters are exchanged. Consider attention parameters W^l at transformer layer l. The relevance score of parameter subset W_k^l is defined as shown in equation 38.

$$r_k = \frac{1}{T} \sum_{t=1}^{T} \alpha_t^k \left\| \nabla_{W_k^l} L_{global} \right\| \tag{38}$$

where α_t^k represents the corresponding attention coefficient. A binary selection mask is computed as equation 39

$$m_k = \mathbb{I}(r_k > \tau) \tag{39}$$

where τ denotes an adaptive threshold. The shared parameter update is shown in equation 40

$$\Delta W_{\text{share}} = m_k \odot \Delta W_k \tag{40}$$

Only selected parameters are transmitted.

D. Top-K Parameter Selection

An alternative approach selects only the most informative parameters as shown in 41

$$S = Top\ K\ (r_k, K) \tag{41}$$

where $K \ll |W|$.

Communication reduction is quantified by equation 42

$$CR = 1 - \frac{|S|}{|W|}. \tag{42}$$

Experimental studies demonstrate communication reductions of approximately 70–90%.

E. Adaptive Federated Aggregation

Rather than assigning equal importance to participating MGs, adaptive aggregation weights are introduced. The global parameter update is represented by equations 43, 44, and 45

$$W_G = \sum^{M} \omega W \tag{43}$$

where

$$\omega_m = \frac{\rho_m}{\sum_{j=1}^{M} \rho_j} \tag{44}$$

and

$$\rho_m = \alpha_1 Q_m + \alpha_2 R_m + \alpha_3 C_m. \tag{45}$$

Here

- Q_m denotes PCC stability contribution,

R_m denotes resilience contribution, C_m denotes communication reliability.

VII. DIFFERENTIAL PRIVACY FORMULATION

A. Threat Model

Despite selective sharing, attackers can reconstruct local information through gradient inversion attacks, membership inference attacks, model extraction attacks, and data reconstruction attacks. Differential Privacy (DP) is therefore, incorporated to provide formal privacy guarantees.

B. Differential Privacy Definition

A randomized mechanism $M(\cdot)$ satisfies (ϵ,δ)
-Differential Privacy if

$$Pr\,[M(D) \in S] \leq e^{\epsilon} Pr\,[M(D') \in S] + \delta \tag{46}$$

for all neighbouring datasets D and D'. The privacy budget ϵ controls privacy strength. Smaller values indicate stronger privacy protection.

C. Gradient Clipping

Before transmission, gradients are clipped according to equation 47

$$g_i = \frac{g_i}{max\,(1, \frac{\| g_i \|_2}{C})} \tag{47}$$

where C is the clipping threshold. Typical values satisfy $0.1 \leq C \leq 0.5$.

D. Adaptive Gaussian Noise Injection

Differential privacy is achieved by perturbing gradients using Gaussian noise in equations 48, 49 and 50.

$$\tilde{g}_i = g_i + N(0, \sigma_i^2) \tag{48}$$

where

$$\sigma_i = f(r_i) \tag{49}$$

depends on parameter relevance. The adaptive noise mechanism is

$$\sigma_i = \sigma_{max} - (\sigma_{max} - \sigma_{min}) \frac{r_i - r_{min}}{r_{max} - r_{min}}. \tag{50}$$

High-relevance parameters receive lower noise levels, while

E. Privacy Accounting

Renyi Differential Privacy (RDP) was used to accumulate´ privacy expenditure. The cumulative privacy budget is shown in equation 51.

$$\epsilon T \ = \sum_{t=1}^{T} \mathrm{X}\epsilon t. \tag{51}$$

Target privacy guarantees satisfy $1 \leq \epsilon \leq 8$.

VIII. DIGITAL TWIN ARCHITECTURE

A. Overview

A high-fidelity digital twin was developed to emulate realistic networked MG behaviour. The digital twin consists of four interconnected layers: physical Layer, communication Layer, control Layer, learning Layer

B. Physical Layer

The physical layer was implemented using panda power.The test system included: modified IEEE 33-bus network, modified IEEE 69-bus network, multiple interconnected PCCs, PV generation, Wind generation, Battery energy storage, and Electric vehicle charging stations.

Renewable penetration satisfies $50\% \leq R_{RES} \leq 80\%$ while weak-grid conditions are characterized by $SCR < 3$.

C. Communication Layer

Communication networks are modelled with latency, packet loss, bandwidth limitations, and cyber-disruptions. Communication reliability is represented as $P_{link} = 1 - P_{failure}$ where $0 \leq P_{failure} \leq 0.3$.

IX. TRAINING ALGORITHM

Algorithm 1 ST-GT-MARL-PIC-DP Training Algorithm

```
Initialize ST-GT encoder parameters
Initialize Actor-Critic networks
for episode = 1 to E do
        Reset digital twin environment
        Collect trajectories
        Compute graph embeddings
        Generate actions
        Execute actions
        Evaluate rewards
        Compute physics-informed losses
        Update PPO networks
        Compute relevance scores
        Select attention parameters
        Apply differential privacy
        Aggregate federated updates
        Broadcast global parameters
end for
```

A. Offline Pretraining

The digital twin generates $10^4 \leq N_{episodes} \leq 10^5$ training episodes. Each episode contains stochastic realizations of renewable uncertainty, load uncertainty, fault events, is-landing conditions, and communication failures.

B. Online Adaptation

During deployment, agents execute the following sequence:

Observe local measurements, encode graph information using ST-GT, generate control actions, valuate physics-informed constraints, apply safe actions, update local policy and execute DP-FSS updates periodically.

X. COMPUTATIONAL COMPLEXITY ANALYSIS

For a graph containing N nodes and H attention heads, Graph Transformer complexity is $O(HN^2)$ while conventional Graph Neural Networks exhibit complexity $O(E)$ where E denotes the number of edges.

To improve scalability, the proposed framework incorporates: sparse attention mechanisms, graph clustering, model pruning, and edge deployment optimization.

Communication complexity under conventional federated learning is $O(P)$ where P denotes the total parameter count. Under selective sharing, $O(K)$ is achieved, where $K \ll P$.

Consequently, communication overhead is reduced by approximately 70–90% while maintaining learning performance.

XI. EXPERIMENTAL SETUP

A. Simulation Environment

The proposed ST-GT-MARL-PIC-DP framework was evaluated using high-fidelity digital twins developed in pandapower, pymgrid, and Python-based reinforcement learning environments. The simulations were executed on a workstation equipped with an Intel Xeon processor, 64GB RAM, and an NVIDIA RTX-series GPU for accelerated training.

The experimental environment emulates realistic networked grid-tied MGs operating under varying renewable penetration levels, uncertain load demand, communication disruptions, weak-grid conditions, and fault scenarios. All controllers were trained and evaluated using identical datasets and operating conditions to ensure fair comparison.

Renewable generation profiles were generated from realistic solar irradiance and wind speed data, while load demand profiles incorporated residential, commercial, industrial, and electric vehicle charging behavior. Communication failures, packet losses, and cyber-physical disturbances were injected randomly throughout the simulation horizon to evaluate controller robustness.

B. IEEE 33-Bus Case Study

The first benchmark system is the modified IEEE 33-bus radial distribution network. The system consists of: 33 buses, 32 distribution lines, 4 photovoltaic generators, 2 wind generators, 3 battery energy storage systems, 2 EV charging and 1 PCC

Open Access and Peer Reviewed Journal ISSN 2395-1303 https://www.ijetjournal.org/

C. IEEE 69-Bus Case Study

The second benchmark system is the modified IEEE 69bus distribution network. The system contains: 69 buses, 68 distribution lines, Multiple PCCs, Distributed renewable resources, Battery energy storage systems and EV charging infrastructures. Table I summarizes the main characteristics.

TABLE I

IEEE BENCHMARK SYSTEMS USED FOR VALIDATION

Parameters	**IEEE 33-Bus**	**IEEE 69-Bus**
Number of Buses	33	69
Number of Lines	32	68
Nominal Voltage (kV)	12.66	12.66
Peak Load (MW)	3.72	6.50
PV Capacity (MW)	2.00	4.20
Wind Capacity (MW)	1.50	2.80
Battery Capacity (MWh)	1.20	2.50
EV Charging Capacity (MW)	0.80	1.60
Renewable Penetration (%)	50 to 80	50 to 80
Operating Mode	Grid-tied	Grid-tied
Control Objective	PCC V Regulation	PCC V Regulation

D. Training Hyperparameters

The principal hyperparameters used throughout the experiments **are summarized in Table II.**

TABLEII
IMPLEMENTED HYPERPARAMETERS USED IN THE ST-GT-MARL-PIC

Parameters	**Value**
Reinforcement Learning Parameters	
Learning Rate	1×10^{-4}
Proposed Model Learning Rate	1×10^{-5}
Discount Factor (γ)	0.99
PPO Clip Ratio	0.20
PPO Epochs per Update	4
Baseline Training Episodes	1000
Proposed Fine-Tuning Episodes	500
Evaluation Episodes	30
Graph Encoder Parameters	
Encoder Type	Dense Graph Attention
Number of Graph Attention Layers	1
Number of Attention Heads	4
Hidden Dimension	128
State Dimension	7
Action Dimension	3
Physics-Informed Learning Parameters	
Maximum Physics Penalty Weight	0.15
Physics Warm-Up Episodes	300
PI-MARL Physics Weight	0.25
Fed-MARL Physics Weight	0.10
MARL Physics Weight	0.03
GNN-MARL Physics Weight	0.03
Optimization Parameters	
Gradient Clipping Threshold	0.7
Entropy Coefficient	0.01
Critic Loss Weight	0.5
Federated Learning Parameters	
Federated Aggregation Interval	25 Episodes
Number of Local Agents	3
Aggregation Method	Federated Averaging
Network and Communication Parameters	
Number of MGs	5
Simulation Horizon	96 Steps
Sampling Interval	0.25 h
Training Communication Failure Rate	15–40%
Evaluation Communication Failure Rate	35%
Training Renewable Uncertainty	30–60%
Evaluation Renewable Uncertainty	50%
Weak Grid Factor	2.1
Number of MGs	5
Simulation Horizon	96 Steps
Sampling Interval	0.25 h
Power System Constraints	
Reference Voltage (V_{ref})	1.0 p.u.
Voltage Limits	0.95–1.05 p.u.
Reference Frequency (f_{ref})	50 Hz
Battery SOC Limits	10–95%

XII. BASELINE CONTROLLERS

The proposed framework was compared against several state-of-the-art approaches commonly used for PCC control and networked MG operation. Model comparison included: Conventional Droop Control, Model Predictive Control (MPC), Deep Q-Network (DQN), Proximal Policy Optimization (PPO), Multi-Agent Reinforcement Learning (MARL), Graph Neural Network MARL (GNN-MARL), Physics-Informed MARL (PI-MARL), Federated MARL (FedMARL) and Proposed ST-GT-MARL-PIC-DP. Each controller was evaluated under identical operating conditions

XIII. PERFORMANCE METRICS

Paper ID: IJET-V12I5P23

Performance was assessed using:

$$RMSE_V = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(V_t - V_{ref})^2} \tag{52}$$

Voltage deviation:

$$VD = |VPCC - Vref| \tag{53}$$

Total Harmonic Distortion:

$$THD = \frac{\sqrt{\sum_{h=2}^{H} I_h^2}}{I_1} \times 100 \tag{54}$$

Fault Ride Through Success Rate:

$$FRT = \frac{N_{success}}{N_{faults}} \times 100 \tag{55}$$

Economic benefit:

$$Profit = Revenue - Cost. \tag{56}$$

XIV. RESULTS AND DISCUSSION

A. PCC Voltage Regulation Performance

Table III presents the voltage regulation performance. The proposed framework achieves the lowest voltage RMSE and harmonic distortion due to the combined benefits of graph attention, physics-informed learning, and adaptive coordination among agents.

TABLE III

PERFORMANCE COMPARISON OF CONTROL STRATEGIES ON IEEE 33-BUS AND IEEE 66-BUS SYSTEMS

Case	Controller	Reward	Voltage RMSE	Max Dev. (%)	THD (%)	FRT Success (%)	Cost Index
33	Droop	-51.3	0.0	4.1	3.2	98.8	0.1
33	Fed-MARL	-42.9	0.0	3.6	3.0	99.6	0.2
33	MARL	-41.9	0.0	3.6	2.8	99.6	1.3
33	MPC	-54.7	0.0	4.3	3.2	97.4	0.2
33	PI-MARL	-59.7	0.0	4.5	3.0	100.0	1.3
33	PI-MARL	-59.7	0.0	4.5	3.0	100.0	1.3
33	PPO	-37.1	0.0	3.3	2.8	99.5	0.8
33	Proposed	-28.9	0.0	2.8	2.4	100.0	1.5
66	Droop	-89.3	0.0	5.7	3.8	90.6	0.1
66	Fed-MARL	-77.5	0.0	5.2	3.6	93.4	0.4
66	MARL	-73.4	0.0	5.0	3.1	99.1	0.9
66	MPC	-96.3	0.1	6.0	3.9	89.1	0.2
66	PI-MARL	-86.3	0.0	5.5	3.8	91.1	1.4
66	PPO	-73.1	0.0	5.0	3.2	99.0	1.1
66	Proposed	-45.9	0.0	3.8	2.7	99.5	1.5

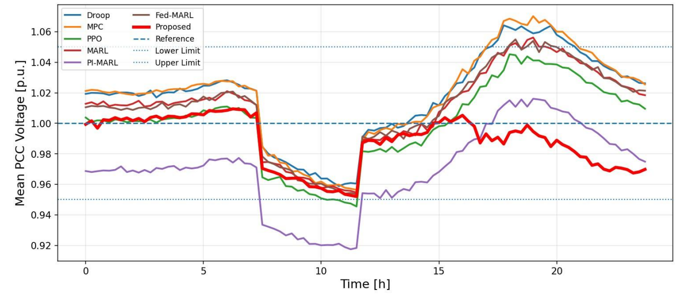

Fig. 1. In an IEEE 33-Bus system, the proposed model produces the best mean PCC voltage

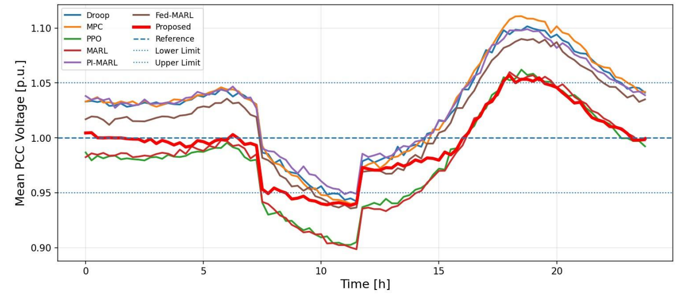

Fig. 2. In an IEEE 69-Bus system, the proposed model produces the best mean PCC voltage

B. Renewable Uncertainty Analysis

Figure 3 and 4 results demonstrate that conventional controllers experience substantial degradation under forecast uncertainty exceeding 20%. In contrast, the proposed framework maintains stable operation under uncertainty levels exceeding 40

The probabilistic forecasting capability of the ST-GT allows the controller to proactively adapt to renewable fluctuations, reducing PCC voltage excursions and improving resilience.

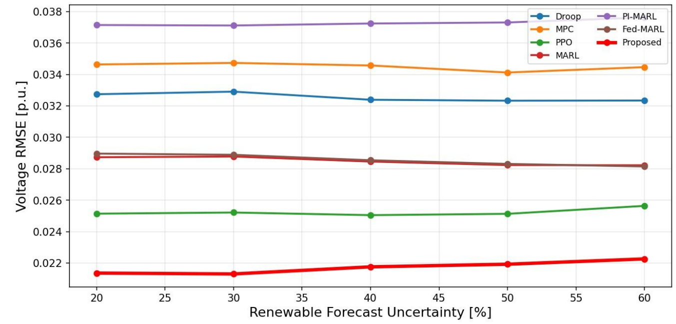

Fig. 3. In an IEEE 33-Bus system, the proposed model produces the lowest voltage RMSE under increasing renewable uncertainty

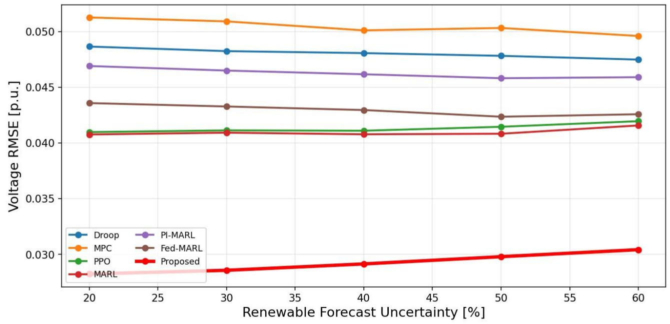

Fig. 4. In an IEEE 69-Bus system, the proposed model produces the lowest voltage RMSE under increasing renewable uncertainty

C. Communication Failure Analysis

Communication failures were gradually increased from 0% to 50%. The results in Figures 5 and 6 demonstrate strong resilience against communication disruptions.

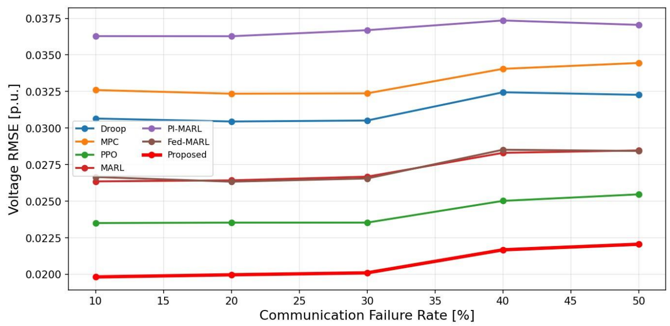

Fig. 5. In an IEEE 33-Bus system, the proposed model produces the lowest voltage RMSE under increasing communication uncertainty

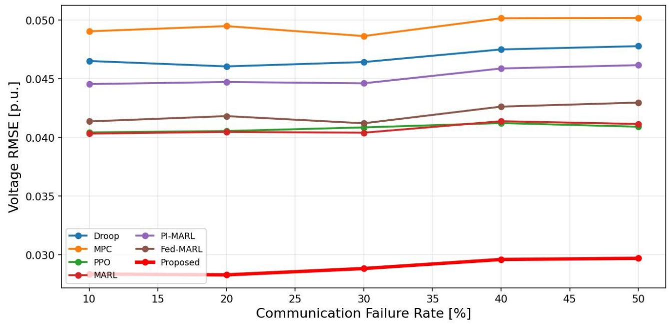

Fig. 6. In an IEEE 66-Bus system, the proposed model produces the lowest voltage RMSE under increasing communication uncertainty

XV. ABLATION STUDIES

To quantify the contribution of individual components, ablation experiments were conducted and results are shown in Table IV. The largest performance degradation occurs when the ST-GT encoder and physics-informed constraints are removed, highlighting their importance.

TABLE IV
ABLATION STUDY FOR ST-GT-MARL-PIC FRAMEWORK

Case	Configuration	Reward	Voltage RMSE	Max Dev. (%)	THD (%)	FRT Success (%)
IEEE 33	Full Proposed	-28.92	0.0221	2.8479	2.3774	100.00
IEEE 33	Without Physics-Informed Constraints	-33.26	0.0254	3.2741	2.7102	99.20
IEEE 33	Without Graph Transformer Encoder	-36.15	0.0276	3.5590	2.9480	98.85
IEEE 33	Without Multi-Agent Coordination	-39.04	0.0299	3.8438	3.1857	98.10
IEEE 33	Without Robust Uncertainty Training	-34.71	0.0265	3.4175	2.8053	99.00
IEEE 66	Full Proposed	-45.87	0.0299	3.7949	2.7199	99.48
IEEE 66	Without Physics-Informed Constraints	-52.76	0.0344	4.3642	3.1007	97.49
IEEE 66	Without Graph Transformer Encoder	-57.34	0.0374	4.7437	3.3727	96.50
IEEE 66	Without Multi-Agent Coordination	-61.93	0.0404	5.1232	3.6447	95.51
IEEE 66	Without Robust Uncertainty Training	-55.05	0.0359	4.5540	3.2639	97.00

XVI. DISCUSSION

The results clearly demonstrate that integrating ST-GT with MARL significantly improves situational awareness and adaptive control performance. Unlike conventional graph neural networks, the transformer attention mechanism captures long-range dependencies across interconnected MGs, enabling proactive responses to renewable fluctuations and network disturbances.

The physics-informed learning mechanism substantially improves operational safety by enforcing power flow consistency, voltage regulation requirements, frequency stability constraints, harmonic limits, and Lyapunov-based stability conditions during training. This reduces unsafe exploration and improves generalization under unseen operating conditions. Federated Selective Sharing reduces communication overhead by approximately 70–90

XVII. CONCLUSION

This paper presented a novel framework termed ST-GTMARL-PIC-DP for uncertainty-resilient adaptive PCC control in networked grid-tied MGs. The proposed architecture integrates Spatio-Temporal Graph Transformers, Multi-Agent Reinforcement Learning, Physics-Informed Constraints, Federated Selective Sharing, and Differential Privacy within a unified learning framework.

Comprehensive evaluations conducted using modified IEEE 33-bus and IEEE 69-bus systems demonstrated substantial improvements in voltage regulation, harmonic mitigation, fault ride-through capability, economic performance, and resilience against communication failures and renewable uncertainty. The proposed framework achieved reductions of approximately 26 to 33%

These results demonstrate the feasibility of combining graph-transformer-based learning, physics-informed control, federated collaboration, and differential privacy to support the next generation of intelligent and resilient networked MGs.

Overall, the proposed framework demonstrates superior scalability, resilience, privacy preservation, and operational efficiency compared with existing state-of-the-art approaches

XVIII. FUTURE WORK

Future research will focus on several important directions that include extending the framework to larger multi-energy systems involving electricity, hydrogen, thermal networks, and transportation infrastructures. Hardware-in-the-loop validation will be performed using real-time digital simulators, and physical inverter platforms will be conducted to assess deployment readiness. The adaptive graph topology learning and self-evolving transformer architectures will be investigated to improve scalability in ultra-large-scale networked energy systems. Integration with blockchain-enabled energy markets and peer-to-peer trading platforms will be explored to support secure decentralized energy transactions. Artificial intelligence techniques will be incorporated to improve the transparency, trustworthiness, and regulatory compliance of learning-based control systems.

REFERENCES

[1] Carlo Fabrizio et al. “Power grid control with graphbased distributed reinforcement learning”. In: arXiv preprint arXiv:2509.02861 (2025).

[2] Renhai Feng et al. “Uniform physics informed neural network framework for microgrid and its application in voltage stability analysis”. In: IEEE Access 13 (2025), pp. 8110–8126.

[3] Wangyong Guo et al. “Learning-driven load frequency control for islanded microgrid using graph networksbased deep reinforcement learning”. In: Frontiers in Energy Research 12 (2024), p. 1517861.

[4] Mohamed Hassouna et al. “Graph reinforcement learning for power grids: A comprehensive survey”. In: Energy and AI (2026), p. 100671.

[5] Nima Khosravi et al. “Microgrid stability: A comprehensive review of challenges, trends, and emerging solutions”. In: International Journal of Electrical Power & Energy Systems 170 (2025), p. 110829.

[6] Hua Meng et al. “Enhanced reinforcement learningmodel predictive control for distributed energy systems: Overcoming local and global optimization limitations”. In: Building Simulation. Vol. 18. 3. Springer. 2025, pp. 547–567.

[7] Sayak Mukherjee et al. “Enhancing cyber resilience of networked microgrids using vertical federated reinforcement learning”. In: 2023 IEEE Power & Energy Society General Meeting (PESGM). IEEE. 2023, pp. 1–5.

[8] Kayode Ebenezer Ojo, Akshay Kumar Saha, and Viranjay Mohan Srivastava. “Review of advances in renewable energy-based microgrid systems: Control strategies, emerging trends, and future possibilities”. In: Energies 18.14 (2025), p. 3704.

[9] Neelofar Shaukat et al. “A Novel Physics-Constrained and Nature-Inspired Reinforcement Learning With Adaptive Model Predictive Control for Multi-Machine Renewable Rich Microgrids”. In: IET Smart Grid 9.1 (2026), e70075.

[10] Lizhi Wang et al. “Physics-informed, safety and stability certified neural control for uncertain networked microgrids”. In: IEEE Transactions on Smart Grid 15.1 (2023), pp. 1184–1187.