Enhancing Cybersecurity and Resilience in Distribution Networks via Hybrid Partitioning and GAN-Driven Dynamic Reconfiguration

Fatemeh Nasr Esfahani, Neeraj Suri, and Xiandong Ma*

***Abstract*—This paper presents a structured hybrid grid partitioning framework designed to enhance cyber-physical resilience and scalability in distribution networks, particularly under high electric vehicle (EV) penetration and evolving cyber threats. The framework integrates three tightly coupled layers. First, a graph-based clustering stage introduces spectral-informed adaptive hierarchical clustering (SIAHC), which combines global spectral features with composite electrical distance metrics to generate modular, self-sufficient, and topologically coherent subnetworks, supporting power loss minimisation, voltage stability, and topological robustness. Second, an optimisation-based refinement layer employs the alternating direction method of multipliers (ADMM) for scalable, distributed coordination across partitions, ensuring feasibility under power flow and voltage constraints. Third, a feedback-informed data-driven layer integrates a Bayesian LSTM variational autoencoder (LSTM-VAE) to forecast cost components and learn clustering parameters, and a conditional Wasserstein GAN (cWGAN-GP) to simulate adversarial scenarios and enable adaptive fallback control under cyber intrusions. The framework is validated on IEEE 33-bus and 123-bus systems with high penetration of EVs, energy storage, and photovoltaics. Results demonstrate improved clustering quality, reduced power loss and voltage deviations, fast convergence, and enhanced cyber-resilience.**

***Index Terms*—Grid Partitioning, Forecasting, Optimisation, Generative Adversarial Networks, Electric Vehicles**

NOMENCLATURE

$\mathcal{V}$, $\mathcal{E}$	Set of all buses and electrical edges in the network
T	Time horizon (number of time steps)
t	Time index
b, b_1, b_2, b_n, b_d	Bus indices
V_b^t, θ_b^t	Voltage magnitude and phase angle at bus b at time t
V_{base}, θ_{base}	System base voltage and phase angle used for normalisation voltage difference terms
V_{nom}	Nominal voltage used for reference in voltage stability cost
$P_{b_1b_2}^t, Q_{b_1b_2}^t$	Active/reactive power flow between buses b_1 and b_2
$P_{bb_n}^t, Q_{bb_n}^t$	Power flow from neighbouring bus b_n to node b
$P_{bb_d}^t$	Active power from node b to downstream bus b_d
P_b^{gen}, P_b^{load}	Active power generation and load at bus b
P_b^{net}	Net power injection at bus b: $P_b^{\text{gen}} - P_b^{\text{load}}$
$R_{b_1b_2}, R_{bb_n}$	Resistance of the transmission line between buses
e, st	EV and ESS indices
E, ST	Total number of EVs and stationary ESSs
x_1, x_2	EV routing location indices
$\Delta x_{x_1,x_2,e}^t$	Routing distance of EV e between locations x_1 and x_2 at time t
$\mathcal{Q}$	Set of all energy storage devices, including ESS and EVs
SoC_{st}^t	State of charge of ESS st at time t
$P_{\text{ch},st}^t, P_{\text{dis},st}^t$	Charging/discharging power of ESS st
$P_{\text{ch},e}^t, P_{\text{dis},e}^t$	Charging/discharging power of EV e
$P_{\text{max},q}$	Maximum power capacity of storage unit $q \in \mathcal{Q}$
r_{EV}^t	Maximum allowable EV routing activity at time t
$\mathcal{K}$, $\mathcal{K}_p$, $\mathcal{K}_{\text{refined}}$	Initial cluster set, p^{th} cluster, and final refined cluster set
C	A specific cluster (or partition)
$\mathcal{P}$, p	Total number of partitions and partition index
$\Gamma_{b_1b_2}$	Binary variable indicating active edge between b_1 and b_2
$w_{b_1b_2}$	Edge importance in partitioning cost $\mathcal{J}_5$
$\mathcal{N}(b)$	Downstream neighbours of bus b
$\mathcal{S}_p^{\text{boundary}}$	Set of boundary buses in partition p used for consensus coordination
$\zeta \in \{V, \theta, P, Q\}$	Set of physical variables involved in consensus updates
$\zeta_{p,\text{boundary}}$	Boundary variable for partition p
$u_{\zeta,p}$	Dual variable for variable ζ in partition p
$\mathcal{J}_1, ..., \mathcal{J}_6$	Cost functions
$\Delta\mathcal{J}_6^t$	Deviation in cybersecurity cost at time t due to intrusion
c_{loss}, c_{volt}, c_{deg}, c_{op}	Cost coefficients for power loss, voltage deviation, battery degradation, and ESS
$\rho_{\text{deg}}, \rho_{\text{op}}$	Battery degradation and operational multipliers
$c_{\text{travel}}, c_{\text{ch}}, c_{\text{dis}}$	EV cost terms for travel, charging, discharging
w_1, w_2, w_5	Objective weights in ADMM
w_a, w_d, w_u	Weights for cybersecurity cost components
$w_{\text{elec}}, w_{\text{sync}}, w_{\text{self}}$	Weights for composite electrical distance
m	Index over spectral weights $\{w_{\text{elec}}, w_{\text{sync}}, w_{\text{self}}\}$
$s_{\text{elec}}, s_{\text{sync}}, s_{\text{self}}$	Spectral scores used to compute distance weights
s_{total}	Sum of all spectral scores for normalisation
$w_{34_1}, w_{34_2}, w_{34_3}$	Weights for cost, posterior regularisation, and clustering feasibility losses in the Bayesian LSTM-VAE
α_θ	Weighting factor adjusting the influence of phase angle misalignment in $\mathcal{J}_5$
β, β_c	Scaling parameters for global and intra-cluster affinity matrix decay
τ_{sw}	Binarisation threshold for relaxed switching variables post-ADMM
$\tau_{\text{var}}, \tau_{\text{entropy}}$	Thresholds for variance and entropy fallback
$\tau_{\text{pri}}, \tau_{\text{dual}}$	Convergence thresholds for ADMM
$\tau_{\text{merging}}, \rho_{\text{penalty}}$	Parameters for AHC merge and modularity preservation
$\tau_{\min}, \tau_{\max}$	Lower/upper bounds on merging threshold
τ_1, τ_2, τ_3	Validation thresholds for clustering quality
τ_c	Minimum size threshold for intra-cluster spectral refinement
τ_{cyber}	Cybersecurity resilience threshold triggering reconfiguration
$\gamma_{\text{penalty}}^k$	Penalty parameter at ADMM iteration k
δ_{penalty}	Gradient penalty coefficient in cWGAN-GP loss
$\mathcal{C}_{\text{phys}}$	Physical plausibility filter for synthetic adversarial vectors
$\widehat{\mathbf{\Theta}}_{\text{cluster}}^t$	Predicted clustering parameters at time t
$\mathbf{\Theta}_{\text{used}}^t$	Final clustering parameters deployed at time t
$\mathbf{\Theta}_{\text{fallback}}^t$	Fallback configuration used under high uncertainty
$\mathbf{\Theta}_{\text{fallback}}^{t,\text{cyber}}$	Cyber-aware fallback clustering parameter set
$\mathbf{\Theta}_{\text{used}}^{t,\text{updated}}$	Cyber-aware fallback clustering parameter set
π_{control}^t	Current control policy at time t
r_p^k, r_d^k	Primal and dual residuals at iteration k
W, L	Affinity matrix and Laplacian
λ_2, λ_3	Fiedler value and third eigenvalue
$\Delta\lambda$	Spectral dispersion across refined clusters ($\lambda_3 - \lambda_2$)
$\vec{v}_2$	Fiedler vector
H_{spec}	Spectral entropy
D^{comp}, D^{elec}, D^{sync}, D^{self}	Composite and component distance metrics
D_{ij}^{eff}	Effective merge distance between clusters C_i and C_j with penalty
$J_{P\theta}$	Jacobian submatrix
$\mathbf{x}_t$	Feature or vector for Bayesian LSTM-VAE
$\mathbf{h}_t$	LSTM hidden state
Ω_t	Latent representation of system state
$\widehat{\mathbf{y}}_t$	Predicted output vector from Bayesian LSTM-VAE
$\mathcal{L}_{34}$	Composite training loss for Bayesian LSTM-VAE
$\mathcal{L}_{\text{penalty}}$	Gradient penalty term enforcing Lipschitz constraint
$\mathcal{L}_{\mathcal{D}_{\text{cyber}}}, \mathcal{L}_{\mathcal{G}_{\text{cyber}}}$	Discriminator and generator losses for the cWGAN-GP
$\mathcal{R}(\widehat{\mathbf{\Theta}}_{\text{cluster}}^t)$	Penalty for infeasible clustering parameter predictions
k	ADMM iteration index
$k_{\max}$	Maximum number of ADMM iterations allowed
N_{gate}	Number of consecutive steps exceeding cyber risk threshold to trigger reconfiguration
ℓ	Temporal window length for LSTM input sequence
$\mathbf{c}_t$	Conditioning vector for cWGAN-GP at time t
$\mathcal{T}_{\text{res}}, \mathcal{T}_{\text{res},\min}, \mathcal{T}_{\text{res},\max}$	Raw, minimum, and maximum system response delays
$\mathbf{v}_{\text{real}}, \mathbf{v}_{\text{fake}}$	Real and generated intrusion vectors
$\mathcal{R}_{\text{cyber}}$	Resilience score
$\mathbf{z}$	Latent noise input to GAN generator
ε	Random scalar for vector interpolation in gradient penalty
ϵ_1, ϵ_2	Small constants for numerical stability and threshold increment
$\varepsilon_{\text{cyber}}$	Small constant in resilience score definition to avoid division by zero
d_v, d_c, d_z, d_{time}, d_π	Dimensions of GAN-related vectors, intrusion, conditioning input, latent noise, temporal encoding, and control policy

*Corresponding author (xiandong.ma@lancaster.ac.uk)

Manuscript received 09/05/2025; revised 08/08/2025; accepted 22/09/2025.

I. Introduction

A. Background and Motivation

THE rapid growth of electric vehicles (EVs) presents major challenges for distribution networks, due to their unpredictable and spatially diverse charging behaviour. This complicates power flow balancing, voltage regulation, and storage scheduling. Also, the increased connectivity of EVs and energy storage systems (ESSs) introduces cyber-physical vulnerabilities like coordinated charging attacks that can disrupt grid stability [1].

Traditional centralised control systems often cannot handle these rapidly changing conditions [2]. Researchers have proposed a range of robust control strategies, such as sliding mode control, H∞ control, and backstepping methods, to enhance resilience and maintain voltage and frequency stability under uncertain conditions (e.g., communication delays, cyber disturbances, and faults) [3], [4]. While effective for local stability and fault tolerance, these strategies typically rely on fixed system topologies and detailed models, thereby limiting scalability and adaptability under dynamic conditions. Multi-agent coordination offers a decentralised alternative, enabling distributed decision-making across agents such as generators, loads, and batteries [5]. This supports plug-and-play operation and fault recovery. However, such methods may be difficult to scale in fast-changing environments and often require high communication overhead. Moreover, the emergent system behaviour can be hard to predict due to decentralised control. Intrusion detection systems have also been deployed to strengthen the cybersecurity of microgrids by detecting anomalies such as false data injection (FDI), denial-of-service (DoS), or topology tampering [6]. These range from rule-based and model-based methods to more advanced machine learning approaches [7]. Although these solutions enhance threat detection, they rely heavily on data quality and tuning, and typically react to threats rather than proactively mitigating vulnerabilities.

In parallel, grid partitioning has emerged as a promising approach to enhance modularity, fault isolation, and resilience by dividing large networks into smaller, flexible clusters that can operate independently or cooperatively during faults or attacks [8], [9]. A wide range of partitioning methods have been proposed, each offering distinct strengths and limitations. Community-based and zone-based clustering are simple and scalable but often overlook critical electrical or operational features unless explicitly integrated [10], [11]. Electrical distance-based methods better capture physical network structure but can be sensitive to topological changes and are computationally intensive [12]. Optimisation-based techniques are goal-driven, targeting load balancing, loss minimisation, or resilience, but tend to be model-dependent and resource-heavy [13]. Spectral and hierarchical clustering provide scalable, structure-aware partitioning using topological properties, though they generally are not adaptable to dynamic conditions unless augmented with data-driven elements such as real-time measurements or historical load profiles [14]. Machine learning-based methods have shown potential for real-time adaptability, but often rely on high-quality labelled data and function as black boxes, limiting physical interpretability [9]. K-means clustering, while fast and straightforward for similarity-based grouping, does not inherently account for network topology or electrical constraints unless these are explicitly encoded [8].

Hence, although each method offers useful capabilities, none fully satisfies the combined requirements of interpretability, adaptability, and cyber-resilience. This motivates the development of hybrid approaches that integrate modelling, optimisation, and learning to support scalable and resilient grid operation under uncertainty.

B. Related Work

Several hybrid partitioning frameworks have been proposed that integrate structural modelling with optimisation or learning. However, few explicitly address cyber-resilience or enable real-time reconfiguration, particularly in distribution systems with high EV penetration. Some approaches incorporate electrical distance metrics alongside optimisation to balance physical realism and operational efficiency. For example, [15] employs Louvain clustering based on voltage sensitivity, followed by flow-based refinement to improve modularity. While structurally effective, real-time adaptability and cyber-physical resilience are not incorporated. A dynamic partitioning approach in [16] focuses on optimally operating a centralised shared energy storage system using multi-objective optimisation and decision-making. It supports real-time adaptability, but it does not offer learning mechanisms or address cyber-resilience and structural interpretability. To support adaptive reconfiguration and cyber-aware operation, [17] proposes a cloud-edge framework that combines deep learning with partitioning rules to respond to abnormal load behaviour. However, it does not explicitly model cyber threats or preserve physical interpretability, limiting its utility in resilience-oriented scenarios.

Several recent hybrid partitioning frameworks incorporate spectral or hierarchical techniques to exploit underlying network structure. For instance, [18] proposes a bi-objective method that combines spectral clustering with weighted consensus to balance active power and voltage controllability. It integrates multiple partitions via consensus strategies, but it does not offer interpretability, real-time adaptability, or cyber-resilience. In [19], a hierarchical spectral clustering approach uses a Laplacian derived from admittance or power flow data and applies unsupervised clustering based on eigenvectors to generate multilevel dendrograms. This method is scalable and interpretable through spectral embedding but assumes static topologies and does not support dynamic operation or cyber-physical robustness. The framework in [20] guides clustering using structural indicators (e.g., electrical coupling and modularity) along with rule-based boundary adjustment, but it remains limited to static scenarios without provisions for cyber-aware adaptation. In [21], spectral graph clustering is used for intentional islanding by minimising inter-island power imbalance with a weighted Laplacian matrix. While operationally sound and structurally grounded, it is not designed for learning-based adaptation or resilience to cyber threats. Spectral feature extraction via singular value decomposition

TABLE I
COMPARISON TABLE: HYBRID PARTITIONING METHODS IN POWER GRID APPLICATIONS

Ref.	Partitioning method	Hybrid components			CP stress testing	Key features and limitations
		Opt.	Learn.	Adapt.		
[16]	Dynamic optimisation-based partitioning for shared energy storage allocation	✓	✗	✓	✗	Optimises storage sharing with adaptive partitioning; no clustering, learning, or cyber-physical aspects.
[17]	Cloud-edge framework combining DL-based operation with partitioning	✓	✓	✓	✗	Deep learning with edge-side rules; adaptable under abnormal loads; no structural interpretability
[18]	Spectral clustering with weighted consensus aggregation for bi-objective partitioning	✓	✗	✗	✗	Balances active power and voltage control; no learning, adaptability, and physical interpretability
[19]	Hierarchical spectral clustering via normalised Laplacian with admittance and power flow weights	✓	✗	✗	✗	Multi-scale clustering preserving internal structure; no adaptive dynamics
[20]	Hierarchical partitioning via structural indicators and modularity heuristics	✓	✗	✗	✗	Incorporates physical structure and boundary adjustment logic; no learning and cyber-physical adaptability
[21]	Spectral graph clustering using weighted Laplacian to minimise power imbalance	✓	✗	✗	✗	Structurally grounded and operationally sound; static design with no learning or adaptive control
[22]	Spectral feature extraction via singular value decomposition with affinity propagation clustering	✓	✗	✗	✗	Combines spectral embedding and unsupervised refinement; no adaptability or cyber-awareness
[23]	Agglomerative hierarchical clustering via full-dimensional electrical distance matrix	✓	✓	✗	✗	Models wind uncertainty and volatility; interpretable via sensitivity metrics; no explicit cyber resilience
This paper	**SIAHC partitioning with ADMM optimisation and LSTM-VAE learning**	✓	✓	✓	✓	**Hybrid and cyber-resilient with structural modelling, learning, and adaptive reconfiguration**

Abbreviations: Opt. = Optimisation; Learn. = Learning-based; Adapt. = Real-time adaptation; CP = Cyber-physical; DL= Deep learning

is combined with affinity propagation clustering and iterative validation in [22]. Despite its unsupervised optimisation and structural grounding, it functions offline and does not offer adaptive or cyber-resilient operation. A reactive voltage partitioning method is introduced in [23] that incorporates wind power uncertainty using agglomerative hierarchical clustering. It enables dynamic reconfiguration and maintains physical interpretability via sensitivity-based metrics by constructing a full-dimensional electrical distance matrix over forecast intervals. However, it does not explicitly address cyber-resilience. Table I provides a comparative summary of these methods.

C. This Work: Contributions and Scope

This paper presents a structured hybrid grid partitioning framework designed to enhance modularity, cyber-physical resilience, and adaptive operation in distribution networks. The framework comprises three complementary components:

First, a novel graph-based clustering method called "spectral-informed adaptive hierarchical clustering (SIAHC)" is introduced. It integrates global spectral properties with composite electrical distance metrics to generate modular, synchronised, and self-sufficient sub-networks. The process involves three stages: (i) spectral pre-analysis for parameter initialisation, (ii) adaptive hierarchical clustering (AHC) guided by physical and operational constraints, and (iii) spectral refinement to enhance intra-cluster resilience and stability.

Second, an optimisation-based refinement stage uses the alternating direction method of multipliers (ADMM) to enforce operational constraints. It decomposes the global problem into parallel subproblems (one per partition) and ensures consistency through consensus on boundary variables. Switching variables are relaxed during iterations and binarised afterward to aid convergence under nonlinear conditions. Unlike mixed-integer programming, ADMM enables scalable, modular optimisation aligned with the pre-defined grid structure [2].

Third, a data-driven adaptation layer improves decision-making under uncertainty by learning from EV mobility, storage dynamics, and cyber threat signals. It integrates two models: (i) a Bayesian LSTM-VAE for forecasting clustering parameters and operational costs based on temporal trends, and (ii) a conditional Wasserstein GAN with gradient penalty (cWGAN-GP) to simulate cyber-physical intrusion scenarios (e.g., FDI, DoS). These models support cyber-resilient operation by informing dynamic reconfiguration and refining clustering weights and fallback policies. Although such architectures have been used for forecasting and anomaly detection, this work uniquely couples a Bayesian LSTM-VAE with an offline-trained cWGAN-GP to actively drive cyber-aware clustering adaptations.

To the best of the authors' knowledge, no existing work integrates graph-based clustering, distributed optimisation, and learning-based resilience testing within a feedback-informed framework that could enhance cyber-physical resilience of EV-intensive distribution systems.

The remainder of this paper is organised as follows: Section II formulates the cost function and operational constraints. Section III details the proposed hybrid grid partitioning framework. Section IV presents case study validations and comparisons with state-of-the-art methods. Section V discusses limitations and directions for future work, and Section VI concludes the paper.

II. COST FUNCTIONS AND CONSTRAINTS

High EV penetration poses several challenges, including increased power losses, voltage instability, complex scheduling, and greater cyber-physical vulnerability. To address these, the proposed framework adopts a modular design with six distinct cost components, each targeting a specific operational goal. Rather than combining them into a single weighted objective, these components are optimised separately across different layers, supporting scalability, modularity, and resilience.

1) Transmission loss cost ($\mathcal{J}_1$): quantifies resistive transmission losses within the distribution network [13]:

$$\mathcal{J}_1 = c_{\text{loss}} \sum_{t=1}^{T} \sum_{\substack{b_1, b_2 \in \mathcal{V} \\ b_1 < b_2}} R_{b_1 b_2} \left[\left(\frac{P_{b_1 b_2}^t}{V_{b_1}^t} \right)^2 + \left(\frac{Q_{b_1 b_2}^t}{V_{b_1}^t} \right)^2 \right], \quad (1)$$

where c_{loss} is the cost coefficient for power loss, T is the time horizon, and $\mathcal{V}$ is the set of network buses. $R_{b_1b_2}$ denotes the resistance between buses b_1 and b_2. The condition $b_1 < b_2$ ensures that each pair of buses is considered only once to avoid double-counting in symmetric configurations. $P^t_{b_1b_2}$, $Q^t_{b_1b_2}$ denote the active and reactive power flows at time t.

2) Voltage deviation penalty ($\mathcal{J}_2$): penalises deviations of bus voltage magnitudes from its nominal value:

$$\mathcal{J}_2 = c_{\text{volt}} \sum_{t=1}^{T} \sum_{b\in\mathcal{V}} \left(\frac{V_b^t - V_{\text{nom}}}{V_{\text{nom}}} \right)^2, \tag{2}$$

where c_{volt} is the cost coefficient for voltage deviations, V_b^t is the voltage at bus b at time t, and V_{nom} is the system's nominal voltage.

3) Stationary ESS cost ($\mathcal{J}_3$): models the costs associated with stationary ESS:

$$\mathcal{J}_3 = \rho_{\text{op}} c_{\text{op}} + \rho_{\text{deg}} c_{\text{deg}} \sum_{st=1}^{ST} \sum_{t=1}^{T} \left(P^t_{\text{ch},st} + P^t_{\text{dis},st} \right), \tag{3}$$

where the total number of ESS units is denoted by ST, with each unit st operating at charging and discharging power levels $P^t_{\text{ch},st}$ and $P^t_{\text{dis},st}$ at time t, respectively. The operational cost includes a scaled fixed term, $\rho_{\text{op}}c_{\text{op}}$, for short-term usage (e.g., maintenance, standby losses), and a degradation cost, $\rho_{\text{deg}}c_{\text{deg}}$, reflecting long-term wear based on energy throughput, where c_{deg} denotes the per-unit battery replacement cost.

4) EV scheduling cost ($\mathcal{J}_4$): models the operational expenses associated with EVs as mobile energy storage units, incorporating travel, charging, and discharging costs across time and space:

$$\mathcal{J}_4 = \sum_{e=1}^{E} \sum_{t=1}^{T} \left(c_{\text{travel}} \Delta x \, X^t_{x_1,x_2,e} + c_{\text{ch}} P^t_{\text{ch},e} + c_{\text{dis}} P^t_{\text{dis},e} \right), \tag{4}$$

where E is the number of EVs, and Δx is the distance between locations x_1 and x_2. The binary variable $X^t_{x_1,x_2,e} \in \{0,1\}$ indicates whether the e^{th} EV travels between these points at time t. The coefficients c_{travel}, c_{ch}, and c_{dis} are the costs per unit distance, charging energy, and discharging energy, respectively. $P^t_{\text{ch},e}$ and $P^t_{\text{dis},e}$ represent the charging and discharging power of the e^{th} EV at time t.

5) Grid partitioning and reconfiguration cost ($\mathcal{J}_5$): promotes efficient and secure grid operation by penalising inconsistencies in voltage and phase angles across reconfigurable transmission lines:

$$\mathcal{J}_5 = \sum_{b_1,b_2} \Gamma_{b_1b_2} w_{b_1b_2} \left(\frac{(V_{b_1} - V_{b_2})^2}{V^2_{\text{base}}} + \alpha_\theta \frac{(\theta_{b_1} - \theta_{b_2})^2}{\theta^2_{\text{base}}} \right), \tag{5}$$

where $\Gamma_{b_1b_2} \in \{0,1\}$ is a binary decision variable indicating whether the link between buses b_1 and b_2 is active, and $w_{b_1b_2}$ reflects the importance of that connection. V_{base} and θ_{base} are normalisation constants for voltage and phase, respectively, while α_θ scales the weight of phase misalignment.

6) Cybersecurity cost ($\mathcal{J}_6$): quantifies the operational risk introduced by cyberattacks and the system's ability to detect and mitigate such threats:

$$\mathcal{J}_6 = w_a(\text{FPR} - \text{ADR}) + w_d \hat{\mathcal{T}}_{\text{res}} + w_u \text{FNR}, \tag{6}$$

where ADR, FPR, and FNR denote the anomaly detection rate, false positive rate, and false negative rate, respectively, all normalised to $[0,1]$ for balanced scaling. These metrics are defined as: $\text{ADR} = \text{TP}/(\text{TP}+\text{FN})$, $\text{FPR} = \text{FP}/(\text{FP}+\text{TN})$, and $\text{FNR} = \text{FN}/(\text{FN}+\text{TP}) = 1 - \text{ADR}$, where TP, FP, TN, and FN denote true positives, false positives, true negatives, and false negatives, respectively. $\hat{\mathcal{T}}_{\text{res}} \in [0,1]$ represents the normalised response delay between attack occurrence and effective mitigation, computed as $\hat{\mathcal{T}}_{\text{res}} = (\mathcal{T}_{\text{res}} - \mathcal{T}_{\text{res,min}})/(\mathcal{T}_{\text{res,max}} - \mathcal{T}_{\text{res,min}})$, where $\mathcal{T}_{\text{res}}$ is the raw measured delay, and $\mathcal{T}_{\text{res,min}}, \mathcal{T}_{\text{res,max}}$ define the expected operating range. The weights w_a, w_d, and w_u prioritise detection accuracy, response speed, and reduction of undetected threats, respectively.

To ensure the stability, safety, and resilience of networked distribution systems with integrated EVs, the proposed optimisation framework incorporates several key operational constraints. These include: (i) power balance, enforcing local active power consistency at each bus via Kirchhoff's Current Law; (ii) voltage limits, keeping bus voltages within safe operating ranges; (iii) line flow limits, preventing thermal overloads in transformers and lines; (iv) storage dynamics, modelling state-of-charge (SoC) evolution for stationary ESSs and EVs; and (v) network radiality, preserving tree-like topologies by prohibiting cycles during cluster merging. These constraints are mathematically formulated as [13]:

$$\begin{cases} P^{\text{gen}}_b + \sum_{q\in\mathcal{Q}_b} P^t_{\text{dis},q} - P^{\text{load}}_b - \sum_{q\in\mathcal{Q}_b} P^t_{\text{ch},q} = \sum_{n\in\mathcal{N}(b)} P^t_{bn}, \quad b \in \mathcal{V}, \\ \sum_{b_n} \left(P^t_{bb_n} - R_{bb_n} \frac{(P^t_{bb_n})^2 + (Q^t_{bb_n})^2}{(V^t_b)^2} \right) = \sum_{b_d\in\mathcal{N}(b)} P^t_{bb_d}, \\ V_{\min} \le V^t_b \le V_{\max}, \quad \forall b \in \mathcal{V}, \\ P^{\min}_{b_1b_2} \le P^t_{b_1b_2} \le P^{\max}_{b_1b_2}, \quad Q^{\min}_{b_1b_2} \le Q^t_{b_1b_2} \le Q^{\max}_{b_1b_2}, \\ \text{SoC}^{t+1}_q = \text{SoC}^t_q + \frac{\eta_{\text{ch},q} P^t_{\text{ch},q}}{c_q} - \frac{P^t_{\text{dis},q}}{\eta_{\text{dis},q} c_q}, \quad \forall q \in \mathcal{Q}, \\ 0 \le \text{SoC}^t_q \le 1, \quad 0 \le \{P^t_{\text{ch},q}, P^t_{\text{dis},q}\} \le P_{\text{max},q}, \quad \forall q \in \mathcal{Q}, \\ |\mathcal{E}_{C_i\cup C_j}| < |\mathcal{V}_{C_i\cup C_j}|, \quad \forall C_i, C_j. \end{cases} \tag{7}$$

where $P^t_{bb_n}$ and $Q^t_{bb_n}$ are the active and reactive power flows between buses at time t, and V^t_b is the voltage magnitude at bus b. Also, P^{gen}_b and P^{load}_b are the generation and load powers, respectively. The set $\mathcal{N}(b)$ includes the downstream neighbours of bus b, and R_{bb_n} is the resistance of the corresponding line. Voltage limits are specified by $V_{\min}$ and $V_{\max}$, while line flow limits are given by $P^{\min\,/\,\max}_{b_1b_2}$ and $Q^{\min\,/\,\max}_{b_1b_2}$. For the set $\mathcal{Q} = \mathcal{Q}_{\text{ESS}} \cup \mathcal{Q}_{\text{EV}}$, which includes both stationary and mobile storage units, the SoC for each device $q \in \mathcal{Q}$ evolves according to efficiency parameters $\eta_{\text{ch},q}$ and $\eta_{\text{dis},q}$, and storage capacity c_q. Charging and discharging power levels, $P^t_{\text{ch},q}$ and $P^t_{\text{dis},q}$, are constrained by the maximum limit $P_{\text{max},q}$. Finally, to maintain a radial network structure, any two candidate clusters C_i and C_j being merged must satisfy $|\mathcal{E}_{C_i\cup C_j}| < |\mathcal{V}_{C_i\cup C_j}|$, where $\mathcal{E}$ and $\mathcal{V}$ are the sets of electrical edges and buses, respectively. It ensures loop-free connectivity and prevents the formation of isolated subnetworks. Notably, two complementary power balance constraints are considered: one ensures nodal energy balance at each bus, while the other accounts for resistive line losses, enhancing the accuracy of voltage drop and power dispatch estimates.

III. Proposed Hybrid Partitioning Framework

To manage the growing complexity of networked distribution systems, driven by high EV and ESS integration and rising cyber-physical threats, a scalable, adaptive, and resilient hybrid partitioning framework is proposed. Unlike traditional static graph- or rule-based methods, this approach integrates three synergistic layers within a feedback-informed architecture.

The process begins with a graph-based clustering layer that partitions the network into modular, self-sufficient subsystems based on electrical coupling, dynamic synchrony, and local self-sufficiency metrics. Intra-cluster spectral analysis is then used to prune weak connections and stabilise voltage profiles, indirectly supporting the minimisation of cost objectives $\mathcal{J}_1$ (power loss), $\mathcal{J}_2$ (voltage stability), and $\mathcal{J}_5$ (reconfiguration cost). Next, an optimisation layer refines each cluster, treating partitions as autonomous agents that minimise local costs while coordinating with neighbouring clusters at shared boundaries. This decentralised strategy preserves network structure and enforces voltage and flow constraints, thereby jointly minimising $\mathcal{J}_1$, $\mathcal{J}_2$, and $\mathcal{J}_5$. To maintain adaptability and resilience under uncertainty, the final layer introduces a data-driven coordination mechanism. It dynamically tunes clustering parameters using forecast models and simulated stress tests, proactively assessing system robustness and triggering fallback reconfiguration when resilience thresholds are breached. This directly targets cost objectives $\mathcal{J}_3$ (stationary ESS), $\mathcal{J}_4$ (EV scheduling), and $\mathcal{J}_6$ (cybersecurity resilience).

Fig. 1 presents the workflow of the proposed hybrid framework. Each stage is broken down into its most critical decision blocks, data exchanges, and feedback mechanisms. For clarity, the flowchart focuses primarily on high-level transitions and does not include lower-level checks, internal flags, or iterative subroutines. These additional implementation details are discussed in the following sections.

A. Step 1: Graph-based clustering

The clustering process follows a three-step SIAHC approach: (i) spectral pre-analysis to extract global structural features and initialise parameters, (ii) constraint-aware adaptive hierarchical clustering (AHC) based on a composite electrical distance metric, and (iii) intra-cluster spectral refinement to split weakly connected clusters and improve modular stability and resilience. Therefore, the proposed SIAHC uses spectral features both as a *prior* to initialise the AHC and as a *posterior* to validate and refine the results to ensure that the clustering is not only physically grounded and modular but also robust against internal fragility. The complete clustering workflow is presented in Algorithm 1. Although spectral analysis precedes hierarchical clustering in implementation, AHC is described first here to clarify its parameterisation, which is subsequently guided by the spectral properties of the Laplacian.

1) Adaptive hierarchical clustering (AHC): The AHC algorithm is guided by a composite electrical distance metric $D^{\text{comp}}_{b_1 b_2}$ which combines static electrical coupling, dynamic synchrony, and local energy self-sufficiency:

$$D^{\text{comp}}_{b_1 b_2} = w_{\text{elec}} \tilde{D}^{\text{elec}}_{b_1 b_2} + w_{\text{sync}} \tilde{D}^{\text{sync}}_{b_1 b_2} + w_{\text{self}} \tilde{D}^{\text{self}}_{b_1 b_2}, \tag{8}$$

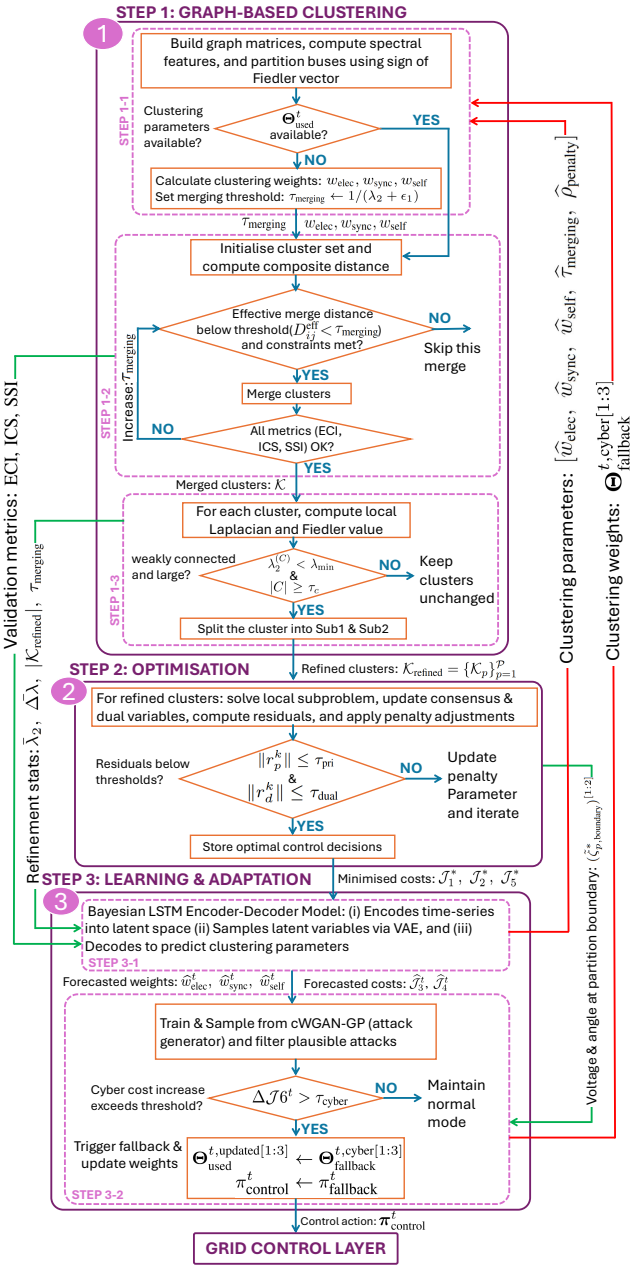


Fig. 1. Proposed hybrid grid partitioning and resilience framework

where w_{elec}, w_{sync}, and w_{self} are weights tuned based on spectral properties of the Laplacian (computed in subsection III-A2).

Let the unnormalised electrical distance between two buses b_1 and b_2 be defined as:

$$D^{\text{elec}}_{b_1 b_2} = J^*_{P\theta}(b_1, b_1) - J^*_{P\theta}(b_1, b_2) - J^*_{P\theta}(b_2, b_1) + J^*_{P\theta}(b_2, b_2), \tag{9}$$

where $J^*_{P\theta}$ is the Moore-Penrose inverse of the active power–phase angle Jacobian submatrix, computed as $J^*_{P\theta} = \left(J^\top_{P\theta} J_{P\theta}\right)^{-1} J^\top_{P\theta}$ [12].

The normalised electrical distance is extracted from:

$$\tilde{D}^{\text{elec}}_{b_1 b_2} = \frac{D^{\text{elec}}_{b_1 b_2} - D^{\text{elec}}_{\min}}{D^{\text{elec}}_{\max} - D^{\text{elec}}_{\min}}, \tag{10}$$

Algorithm 1 SIAHC for graph-based partitioning

Inputs: $(\mathcal{V}, \mathcal{E})$; constraints; ρ_{penalty}; τ_c; $\mathbf{\Theta}^t_{\text{used}}$ from *Step 3* (Section III-C)
Output: $\mathcal{K}_{\text{refined}} = \{\mathcal{K}_1, \ldots, \mathcal{K}_{\mathcal{P}}\}$, each with boundary set $\mathcal{S}_p^{\text{boundary}}$

Step 1-1: Spectral pre-analysis

Compute $\mathbf{W}, \mathbf{L}, \lambda_2, \lambda_3, \vec{v}_2, \Delta\lambda, H_{\text{spec}}$
Partition $\mathcal{V}$ into Groups A and B using the sign of $\vec{v}_2$
if $\mathbf{\Theta}^t_{\text{used}}$ is available **then**
  $(w_{\text{elec}}, w_{\text{sync}}, w_{\text{self}}, \tau_{\text{merging}}, \rho_{\text{penalty}}) \leftarrow$ values from $\mathbf{\Theta}^t_{\text{used}}$
else
  Compute weights $w_{\text{elec}}, w_{\text{sync}}, w_{\text{self}}$
  Set merging threshold: $\tau_{\text{merging}} \leftarrow 1/(\lambda_2 + \epsilon_1)$
end if

Step 1-2: Adaptive hierarchical clustering

Initialise $\mathcal{K} \leftarrow \{C_b\}_{b \in \mathcal{V}}$
Compute $D^{\text{comp}}_{b_1 b_2}$ (using current weights)
`valid` $\leftarrow$ `False`;
Enable options: `greedy` $\leftarrow$ `True`, `use_ANN` $\leftarrow$ `True`
while `not valid` **do**
  `merged` $\leftarrow$ `True`
  while `merged` **do**
    `merged` $\leftarrow$ `False`
    Identify pair (C_i, C_j) with minimum D^{comp}_{ij}
    if C_i, C_j belong to opposite Fiedler groups **then**
      $D^{\text{eff}}_{ij} \leftarrow D^{\text{comp}}_{ij} + \rho_{\text{penalty}_{ij}}$
    else
      $D^{\text{eff}}_{ij} \leftarrow D^{\text{comp}}_{ij}$
    end if
    if $D^{\text{eff}}_{ij} < \tau_{\text{merging}}$ and constraints satisfied **then**
      Merge: $C_i \leftarrow C_i \cup C_j$, update $\mathcal{K}$,
      `merged` $\leftarrow$ `True`
    else
      `merged` $\leftarrow$ `False`
    end if
  end while
  Evaluate clustering using ECI, ICS, SSI
  if All metrics satisfy thresholds **then**
    `valid` $\leftarrow$ `True`
  else
    if `greedy` **then**
      Preserve validated clusters and incrementally refine unstable ones
    else
      Reset: $\mathcal{K} \leftarrow \{C_b \mid b \in \mathcal{V}\}$
      Re-compute $D^{\text{comp}}_{b_1 b_2}$ if weights changed
    end if
    Update merging threshold: $\tau_{\text{merging}} \leftarrow \tau_{\text{merging}} + \epsilon_2$
  end if
end while

Step 1-3: Intra-cluster spectral refinement

$\mathcal{K}_{\text{refined}} \leftarrow \emptyset$
for each cluster $C \in \mathcal{K}$ **do**
  Construct local matrices $\mathbf{W}_C, \mathbf{L}_C$
  Extract $\lambda_2^{(C)}$ and Fiedler vector $\vec{v}_2^{(C)}$
  `refined` $\leftarrow$ `False`
  if $\lambda_2^{(C)} < \lambda_{\min}$ and $|C| \geq \tau_c$ **then**
    Partition C into Subgroups Sub_1, Sub_2 via sign of $\vec{v}_2^{(C)}$
    Compute new values $\lambda_2^{(\text{Sub}_1)}$, $\lambda_2^{(\text{Sub}_2)}$
    if Both improved **and** constraints satisfied **then**
      Add Sub_1, Sub_2 to $\mathcal{K}_{\text{refined}}$
      `refined` $\leftarrow$ `True`
    end if
  end if
  if `not refined` **then**
    Add C to $\mathcal{K}_{\text{refined}}$
  end if
end for
return $\mathcal{K}_{\text{refined}} \longrightarrow$ *Step 2* (Section III-B);
ECI_t, ICS_t, SSI_t, $\bar{\lambda}_2$, $\Delta\bar{\lambda}$, $|\mathcal{K}_{\text{refined}}|$, $\tau_{\text{merging}} \longrightarrow$ *Step 3-1* (Section III-C1)

where $D^{\text{elec}}_{\min}$ and $D^{\text{elec}}_{\max}$ are the minimum and maximum values of $D^{\text{elec}}_{b_1 b_2}$ computed across all bus pairs. Similarly, the normalised synchronisation distance is computed as:

$$\tilde{D}^{\text{sync}}_{b_1 b_2} = \frac{1 - \frac{\text{Cov}(\theta_{b_1}(t), \theta_{b_2}(t))}{\sigma_{b_1}\sigma_{b_2}} - D^{\text{sync}}_{\min}}{D^{\text{sync}}_{\max} - D^{\text{sync}}_{\min}}, \tag{11}$$

where $\text{Cov}(\theta_{b_1}(t), \theta_{b_2}(t))$ is the covariance of voltage angle time series and $\sigma_{b_1}, \sigma_{b_2}$ are their respective standard deviations. Finally, the normalised self-sufficiency distance $\tilde{D}^{\text{self}}_{b_1 b_2}$ is computed as:

$$\tilde{D}^{\text{self}}_{b_1 b_2} = \frac{\left| \left(\frac{P^{\text{gen}}_{b_1} - P^{\text{load}}_{b_1}}{P^{\text{gen}}_{b_1} + P^{\text{load}}_{b_1}} \right) - \left(\frac{P^{\text{gen}}_{b_2} - P^{\text{load}}_{b_2}}{P^{\text{gen}}_{b_2} + P^{\text{load}}_{b_2}} \right) \right| - D^{\text{self}}_{\min}}{D^{\text{self}}_{\max} - D^{\text{self}}_{\min}}, \tag{12}$$

Given a candidate clustering $\mathcal{K}$, the clustering objective minimises the intra-cluster composite distances:

$$\min_{\mathcal{K}} \sum_{C \in \mathcal{K}} \sum_{\substack{b_1, b_2 \in C \\ b_1 < b_2}} D^{\text{comp}}_{b_1 b_2}, \tag{13}$$

Aside from improving modularity, minimising intra-cluster composite distances enhances cyber-resilience by creating compact, well-isolated clusters. This helps contain disruptions locally and reduce inter-cluster exposure.

As shown in *Step 1-2* of Algorithm 1, the AHC procedure proceeds through five main stages: initialisation, modularity-aware merging, iterative merging, validation, and termination. During initialisation, each bus $b \in \mathcal{V}$ is treated as an individual cluster C_b, and a composite electrical distance matrix $D^{\text{comp}}_{b_1 b_2}$ is computed using three spectral-informed weights: electrical proximity (w_{elec}), dynamic synchrony (w_{sync}), and local self-sufficiency (w_{self}). These weights, along with the initial merging threshold $\tau^{(0)}_{\text{merging}}$, are derived from spectral pre-analysis in *Step 1-1*. To discourage premature or physically implausible merges, a modularity-preserving penalty ρ_{penalty} is applied if candidate clusters originate from opposite sides of the Fiedler split. In such cases, the effective merge distance is computed as $D^{\text{eff}}_{ij} = D^{\text{comp}}_{ij} + \rho_{\text{penalty}_{ij}}$.

In each iteration, the algorithm merges the pair of clusters with the lowest effective distance D^{eff}_{ij}, provided it is below the threshold τ_{merging} and the resulting subgraph remains sparse. An optional `use_ANN` flag accelerates merging via approximate nearest neighbour search.

Following merging, validation is performed using three metrics: (1) electrical cohesiveness index (ECI), which measures intra-cluster compactness, (2) inter-cluster separation (ICS), which assesses distinctiveness between clusters, and (3) self-sufficiency index (SSI), which evaluates energy autonomy within each cluster:

$$\begin{cases} \text{ECI} = 1 - \dfrac{\sum\limits_{b_1 \in \mathcal{V}} \sum\limits_{b_2 \in \mathcal{N}(b_1)} D^{\text{comp}}_{b_1 b_2}}{\sum\limits_{b_1, b_2 \in \mathcal{V}} D^{\text{comp}}_{b_1 b_2}}, \text{ICS} = \dfrac{\sum\limits_{b_1 \in \mathcal{V}} \sum\limits_{b_2 \notin \mathcal{N}(b_1)} D^{\text{comp}}_{b_1 b_2}}{\sum\limits_{b_1, b_2 \in \mathcal{V}} D^{\text{comp}}_{b_1 b_2}}, \\ \text{SSI} = 1 - \dfrac{\left| \sum\limits_{b \in C} P^{\text{net}}_b \right|}{\sum\limits_{b \in C} \left(|P^{\text{gen}}_b| + |P^{\text{load}}_b| \right)}. \end{cases} \tag{14}$$

Acceptable clusters must satisfy the thresholds ECI $\geq \tau_1$, ICS $\geq \tau_2$, and SSI $\geq \tau_3$. If a cluster fails validation, the merging threshold τ_{merging} is adjusted, and cluster assignments may be revised using nearest-neighbour reassignment. Meanwhile, the greedy termination logic governs the algorithm's response to persistent validation failure. If the `greedy` flag is disabled, the algorithm resets all clusters to singletons and restarts the process. If enabled, it retains validated clusters and continues refining only the unstable ones.

2) Spectral pre-analysis and parameter initialisation: Before hierarchical clustering begins, a spectral pre-analysis is performed to extract global structural features from the network graph and initialise the key parameters for AHC (i.e., the distance weights ($w_{\text{elec}}, w_{\text{sync}}, w_{\text{self}}$) and the adaptive merging threshold τ_{merging}).

The distribution network is represented as a weighted undirected graph $(\mathcal{V}, \mathcal{E})$. Each edge is weighted by electrical proximity as $W_{b_1 b_2} = \exp\left(-\frac{D^{\text{comp}}_{b_1 b_2}}{\beta^2}\right)$, where $W_{b_1 b_2}$ is the affinity between buses b_1 and b_2, and β is a positive scaling parameter that controls the decay of similarity with distance. The corresponding Laplacian $\mathbf{L} = \mathbf{D} - \mathbf{W}$ is constructed, and its spectral decomposition yields three critical indicators: (i) the Fiedler value λ_2 which represents algebraic connectivity (ii) the spectral gap $\Delta\lambda = \lambda_3 - \lambda_2$, indicating modular separation, and (iii) the spectral entropy H_{spec}, capturing the uniformity of energy distribution in the spectrum.

A preliminary partition is obtained via the sign of the Fiedler vector $\vec{v}_2$ to separate buses into two groups as: $\text{A} = \{b \in \mathcal{V} \mid \vec{v}_2(b) \geq 0\}$ and $\text{B} = \{b \in \mathcal{V} \mid \vec{v}_2(b) < 0\}$.

These spectral quantities inform the initial parameter settings for AHC via the following qualitative mappings:

$$\begin{cases} \Delta\lambda \text{ large} & \Rightarrow \uparrow w_{\text{elec}} \quad \text{(emphasise modularity)} \\ \lambda_2 \text{ low} & \Rightarrow \downarrow \tau_{\text{merging}} \quad \text{(limit premature merges)} \\ \lambda_2 \text{ low} & \Rightarrow \uparrow w_{\text{sync}} \quad \text{(emphasise synchrony)} \\ H_{\text{spec}} \text{ high} & \Rightarrow \uparrow w_{\text{self}} \quad \text{(promote self-sufficiency)} \end{cases}$$

To formalise these heuristics, let us define:

$$s_{\text{elec}} = \Delta\lambda, \quad s_{\text{sync}} = \frac{1}{\lambda_2 + \epsilon_1}, \quad s_{\text{self}} = H_{\text{spec}}, \tag{15}$$

which are normalised to obtain weights $w_m = s_m / s_{\text{total}}$, where $m \in \{\text{elec}, \text{sync}, \text{self}\}$ and $s_{\text{total}} = s_{\text{elec}} + s_{\text{sync}} + s_{\text{self}}$.

The spectral entropy is computed as $H_{\text{spec}} = -\sum_{\nu=2}^{|\mathcal{V}|} \tilde{\lambda}_\nu \log \tilde{\lambda}_\nu$, where $\tilde{\lambda}_\nu$ denotes the normalised eigenvalue magnitude. The initial merging threshold is also derived from spectral properties as $\tau^{(0)}_{\text{merging}} = 1/(\lambda_2 + \epsilon_1)$, where ϵ_1 is a small regularisation constant to avoid division by zero. It should be noted that these parameters guide the initial clustering behaviour, but they are not fixed and are adaptively updated in *Step 3*.

3) Spectral refinement for intra-cluster stability: After initial clustering via AHC, a spectral refinement stage is applied to strengthen intra-cluster connectivity and modular stability. Each cluster $C \in \mathcal{K}$ is treated as a subgraph with its own Laplacian $\mathbf{L}_C = \mathbf{D}_C - \mathbf{W}_C$, where the affinity matrix $\mathbf{W}_C$ is computed using intra-cluster composite electrical distances, scaled by a local decay parameter β_c as $W^{(C)}_{b_1 b_2} = \exp\left(-\frac{D^{\text{comp}}_{b_1 b_2}}{\beta_c^2}\right)$.

The cluster's algebraic connectivity is assessed via its Fiedler value $\lambda_2^{(C)}$. If $\lambda_2^{(C)} < \lambda_{\min}$ and the cluster size satisfies $|C| \geq \tau_c$, the cluster is bisected into two subgroups based on the sign of $\vec{v}_2^{(C)}$:

$$\begin{cases} \text{Subgroup Sub}_1 = \left\{b \in C \mid \vec{v}_2^{(C)}(b) \geq 0\right\}, \\ \text{Subgroup Sub}_2 = \left\{b \in C \mid \vec{v}_2^{(C)}(b) < 0\right\} \end{cases} \tag{16}$$

Algorithm 2 ADMM-based operational refinement

Inputs: Refined partitions $\mathcal{K}_{\text{refined}} = \{\mathcal{K}_p\}_{p=1}^{\mathcal{P}}$; initial normalised state variables $(\tilde{\zeta}^0_{p,\text{boundary}}, \Gamma^0_p)$; initial penalty parameter $\gamma^0_{\text{penalty}}$; dual variables $\{\bar{u}^0_{\zeta,p}\}_{\zeta \in \{V,\theta,P,Q\}}$; convergence thresholds $\tau_{\text{pri}}, \tau_{\text{dual}}$; maximum iterations $k_{\max}$; threshold $\tau_{\text{sw}} \in (0,1)$
Output: Final operational states $(\zeta^*_p,\ \Gamma^{\text{bin}}_p,\ \tilde{\zeta}^*_{p,\text{boundary}})$ for all $\mathcal{K}_p \in \mathcal{K}_{\text{refined}}$; cost values $(\mathcal{J}_1^*,\ \mathcal{J}_2^*,\ \mathcal{J}_5^*)$
`converged` $\leftarrow$ `False`
$k \leftarrow 0$ ▷ Initialise iteration counter
while `not converged` **do**
  for all $\mathcal{K}_p \in \mathcal{K}_{\text{refined}}$ **(in parallel) do**
    Solve relaxed local subproblem for continuous states $(V_p^{k+1}, \theta_p^{k+1}, P_p^{k+1}, Q_p^{k+1})$ and relaxed switching variables $\Gamma_p^{k+1} \in [0,1]$ using Eq. (19)
  end for
  for all $\zeta \in \{V, \theta, P, Q\}$ and $p = 1$ to $\mathcal{P}$ **do**
    Update consensus variables and dual variables using Eq. (20)
  end for
  Compute primal and dual residuals (r_p^k, r_d^k)
  Update penalty parameter $\gamma^k_{\text{penalty}}$
  if $\|r_p^k\| \leq \tau_{\text{pri}}$ **and** $\|r_d^k\| \leq \tau_{\text{dual}}$ **then**
    `converged` $\leftarrow$ `True`
  else if $k \geq k_{\max}$ **then**
    `converged` $\leftarrow$ `True` ▷ Force termination (non-converged)
  else
    $k \leftarrow k + 1$
  end if
end while
for all $\mathcal{K}_p \in \mathcal{K}_{\text{refined}}$ **do**
  Project relaxed switching states: $\Gamma^{\text{bin}}_p = \mathbb{I}[\Gamma^*_p \geq \tau_{\text{sw}}]$
end for
return $(\tilde{\zeta}^*_{p,\text{boundary}})^{[1:2]} \longrightarrow$ *Step 3-2* (Section III-C2);
  $(\mathcal{J}_1^*,\ \mathcal{J}_2^*,\ \mathcal{J}_5^*) \longrightarrow$ *Step 3-1* (Section III-C1)

The split is accepted only if both resulting sub-clusters exhibit improved algebraic connectivity and remain compliant with operational constraints. Clusters smaller than the threshold τ_c are exempt from refinement unless they violate these constraints.

The main output of this step is the refined cluster set $\mathcal{K}_{\text{refined}} = \{\mathcal{K}_p\}_{p=1}^{\mathcal{P}}$. For each refined cluster $\mathcal{K}_p \in \mathcal{K}_{\text{refined}}$, the set of boundary buses is defined as:

$$\mathcal{S}_p^{\text{boundary}} \leftarrow \{b_1 \in \mathcal{K}_p \mid \exists\, b_2 \notin \mathcal{K}_p \text{ such that } (b_1, b_2) \in \mathcal{E}\}, \tag{17}$$

i.e., the buses in $\mathcal{K}_p$ that are directly connected to nodes in other clusters. These sets are finalised and used in *Step 2* to enforce inter-cluster consensus constraints during distributed optimisation.

B. Step 2: Optimisation-based operational refinement

This step focuses on optimising operational performance by performing distributed optimisation over the refined partitions $\mathcal{K}_{\text{refined}}$ generated in *Step 1*. The global coordination objective aggregates partition-level costs across the network as:

$$\min_{\{V_p^t, \theta_p^t, P_p^t, Q_p^t, \Gamma_p\}} \sum_{p=1}^{\mathcal{P}} \left(w_1 \mathcal{J}_{1,p} + w_2 \mathcal{J}_{2,p} + w_5 \mathcal{J}_{5,p}\right), \tag{18}$$

subject to several operational constraints (detailed in section II). The weights w_1, w_2, w_5 serve as empirically tunable multipliers used to adjust the importance of each objective.

Each partition $\mathcal{K}_{\text{refined}} = \{\mathcal{K}_p\}_{p=1}^{\mathcal{P}}$ operates as an independent optimisation agent. The coordination problem is solved using an ADMM-based algorithm (Algorithm 2), which allows parallel solving of local subproblems while maintaining consistency at shared boundaries. Its decision variables include

voltage magnitudes and phase angles (V_p^t, θ_p^t), active/reactive powers (P_p^t, Q_p^t), and switching states Γ_p:

$$\begin{aligned}(V_p^{k+1}, \theta_p^{k+1}, P_p^{k+1}, Q_p^{k+1}, \Gamma_p^{k+1}) = \\ \arg \min_{V_p, \theta_p, P_p, Q_p, \Gamma_p \in [0,1]} \Bigg[w_1 \mathcal{J}_{1,p} + w_2 \mathcal{J}_{2,p} + w_5 \mathcal{J}_{5,p} \\ + \frac{\gamma_{\text{penalty}}^k}{2} \sum_{\zeta \in \{V,\theta,P,Q\}} \left\| \tilde{\zeta}_{p,\text{boundary}} - \tilde{\bar{\zeta}}^k + \tilde{u}_\zeta^k \right\|^2 \Bigg],\end{aligned} \tag{19}$$

where $\tilde{\zeta} = \zeta / \zeta_{\text{base}}$ denotes the normalised form of each electrical variable $\zeta \in \{V, \theta, P, Q\}$. Coordination across clusters is enforced at the boundary buses $\mathcal{S}_p^{\text{boundary}}$, which connect each refined partition $\mathcal{K}_p$ to neighbouring clusters. At these interfaces, consensus is achieved by enforcing $\zeta_{p,\text{boundary}} = \bar{\zeta}$, ensuring inter-cluster consistency of shared state variables. The variable penalty parameter $\gamma_{\text{penalty}}^k$ is updated adaptively at each iteration to balance convergence between primal and dual residuals (r_p^k, r_d^k). If the primal residual dominates, $\gamma_{\text{penalty}}^k$ is increased, and if the dual residual is larger, it is decreased. When the two are balanced within a given threshold, the penalty remains unchanged.

After solving the local subproblems, consensus and dual variables are updated via averaging:

$$\begin{cases} \tilde{\bar{\zeta}}^{k+1} = \dfrac{1}{\mathcal{P}} \displaystyle\sum_{p=1}^{\mathcal{P}} \left(\tilde{\zeta}_{p,\text{boundary}}^{k+1} + \tilde{u}_{\zeta,p}^k \right), \\ \tilde{u}_{\zeta,p}^{k+1} = \tilde{u}_{\zeta,p}^k + \left(\tilde{\zeta}_{p,\text{boundary}}^{k+1} - \tilde{\bar{\zeta}}^{k+1} \right). \end{cases} \tag{20}$$

The procedure iterates until residuals fall below predefined thresholds or maximum iterations are reached. After convergence, the relaxed switching decisions $\Gamma_p^* \in [0,1]$ are projected to binary decisions via a thresholding operation as: $\Gamma_p^{\text{bin}} = \mathbb{I}[\Gamma_p^* \geq \tau_{\text{sw}}]$, where $\tau_{\text{sw}} \in (0,1)$ and $\mathbb{I}[\cdot]$ denotes the indicator function, returning 1 if the condition inside the brackets is true.

Each cluster returns its final operational states ζ_p^*, binary switching decisions Γ_p^{bin}, and consensus-aligned boundary variables $\tilde{\zeta}_{p,\text{boundary}}^*$. These are used to compute the cost components $\mathcal{J}_1^*, \mathcal{J}_2^*, \mathcal{J}_5^*$, which serve as inputs to *Step 3-1* for forecasting. Meanwhile, the first two components of the boundary variables, $(\tilde{\zeta}_{p,\text{boundary}}^*)^{[1:2]}$, are passed to *Step 3-2* to support resilience assessment under adversarial stress testing.

C. Step 3: Data-driven learning and adaptive reconfiguration for cyber-physical resilience

Steps 1 and *2* of the proposed framework generate structurally modular and operationally optimised partitions based on physics-aware clustering (via SIAHC) and distributed optimisation (via ADMM). However, these steps are based on statically tuned parameters (i.e., $w_{\text{elec}}, w_{\text{sync}}, w_{\text{self}}, \tau_{\text{merging}}$, and ρ_{penalty}), which may not remain optimal under dynamic grid conditions. Moreover, some operational costs (e.g., $\mathcal{J}_3$ for storage, $\mathcal{J}_4$ for EV scheduling, and $\mathcal{J}_6$ for cyber resilience) cannot be easily modelled with static optimisation.

To address these limitations, *Step 3* develops a data-driven coordination layer. This layer adaptively updates clustering parameters using historical data as well as learnt cost forecasts and simulated cyber-physical disruptions. *Step 3* consists of two interconnected modules:

1) Step 3-1: Data-driven learning of clustering parameters: This step forecasts storage and mobility costs $\mathcal{J}_3$ and $\mathcal{J}_4$ using a Bayesian LSTM-VAE trained on historical data. Based on these forecasts, it predicts clustering parameters (i.e., $w_{\text{elec}}, w_{\text{sync}}, w_{\text{self}}, \tau_{\text{merging}}$, and ρ_{penalty}). These parameters are then used in *Step 1*.

The input vector at time t captures the dynamic relationship between storage behaviour, EV mobility, current clustering configuration, and recent cost values:

$$\begin{aligned}\mathbf{x}_t = [& \underbrace{\sum_q \text{SoC}_q^t,\ \sum_q P_{\text{ch},q}^t,\ \sum_q P_{\text{dis},q}^t}_{\text{EV/ESS}}, \ \underbrace{\sum_{x_1,x_2,e} \Delta x_{x_1,x_2,e}^t X_{x_1,x_2,e}^t}_{\text{EV routing distance}}, \\ & \underbrace{\mathcal{J}_1^*,\ \mathcal{J}_2^*,\ \mathcal{J}_5^*}_{\text{Lagged costs (Step 2)}}, \ \underbrace{\text{ECI}_t,\ \text{ICS}_t,\ \text{SSI}_t,\ \bar{\lambda}_2,\ \bar{\Delta\lambda},\ |\mathcal{K}_{\text{refined}}|,\ \tau_{\text{merging}}}_{\text{Clustering parameters (Step 1)}}],\end{aligned} \tag{21}$$

where $\bar{\lambda}_2 = \frac{1}{|\mathcal{K}_{\text{refined}}|} \sum_{C \in \mathcal{K}_{\text{refined}}} \lambda_2^{(C)}$ is the average Fiedler value and $\bar{\Delta\lambda} = \frac{1}{|\mathcal{K}_{\text{refined}}|} \sum_{C \in \mathcal{K}_{\text{refined}}} \left(\lambda_3^{(C)} - \lambda_2^{(C)} \right)$. As shown, EV and ESS operations are also taken into account, influencing the stationary cost $\mathcal{J}_3$ and EV cost $\mathcal{J}_4$.

The input sequence $\{\mathbf{x}_{t,1}, \ldots, \mathbf{x}_{t,\ell}\}$, defined over a temporal window of length ℓ, is processed by a Bayesian LSTM encoder to model temporal dependencies and encode grid behaviour with uncertainty awareness. To account for posterior uncertainty, multiple latent samples are drawn via Monte Carlo sampling and passed through the decoder. The final prediction is obtained by averaging the decoded outputs over all samples and consists of two parallel multi-layer networks:

$$\hat{\mathbf{y}}_t = \Bigg[\underbrace{\widehat{w}_{\text{elec}}^t,\ \widehat{w}_{\text{sync}}^t,\ \widehat{w}_{\text{self}}^t,\ \widehat{\tau}_{\text{merging}}^t,\ \widehat{\rho_{\text{penalty}}}^t}_{\hat{\mathbf{y}}_t^{(1)} = \widehat{\boldsymbol{\Theta}}_{\text{cluster}}^t}, \ \underbrace{\widehat{\mathcal{J}}_3^t,\ \widehat{\mathcal{J}}_4^t}_{\hat{\mathbf{y}}_t^{(2)}} \Bigg] \in \mathbb{R}^7 \tag{22}$$

The Bayesian LSTM-VAE is trained using the composite objective $\mathcal{L}_{34}$ that balances cost prediction accuracy, posterior regularisation, and clustering feasibility:

$$\begin{aligned}\mathcal{L}_{34} = \ & w_{34_1} \left(\left| \frac{\widehat{\mathcal{J}}_3^t - \mathcal{J}_3^t}{\bar{\mathcal{J}}_3} \right| + \left| \frac{\widehat{\mathcal{J}}_4^t - \mathcal{J}_4^t}{\bar{\mathcal{J}}_4} \right| \right) \\ & + w_{34_2} \text{KL}\left(q_\phi(\Omega_t \mid \mathbf{h}_t) \,\|\, \mathcal{N}(0, I) \right) + \ w_{34_3} \cdot \mathcal{R}(\widehat{\boldsymbol{\Theta}}_{\text{cluster}}^t)\end{aligned} \tag{23}$$

where w_{34_1}, w_{34_2}, and w_{34_3} are weights that control the trade-off between cost prediction accuracy, latent space regularisation, and feasibility of the clustering parameters. Normalisation is applied over the training dataset $\mathcal{D}_T$ (e.g., $\bar{\mathcal{J}}_3 = \frac{1}{|\mathcal{D}_T|} \sum_t \mathcal{J}_3^t$). The second term adds a Kullback–Leibler (KL) divergence between the learnt latent posterior $q_\phi(\Omega_t \mid \mathbf{h}_t)$, which models the distribution of latent variables Ω_t given the encoded input history $\mathbf{h}_t$, and a standard normal prior $\mathcal{N}(0, I)$. This acts as a regulariser and helps the model learn structured, uncertainty-aware representations. The third term $\mathcal{R}(\widehat{\boldsymbol{\Theta}}_{\text{cluster}}^t)$ applies a soft penalty on the predicted clustering parameters:

$$\begin{aligned}& \widehat{w}_{\text{elect}}^t,\ \widehat{w}_{\text{sync}}^t,\ \widehat{w}_{\text{self}}^t \geq 0, \quad \widehat{w}_{\text{elect}}^t + \widehat{w}_{\text{sync}}^t + \widehat{w}_{\text{self}}^t \leq 1, \\ & \tau_{\min} \leq \widehat{\tau}_{\text{merging}}^t \leq \tau_{\max}, \quad 0 < \widehat{\rho_{\text{penalty}}}^t \leq \rho_{\text{penalty}}^{\max}.\end{aligned} \tag{24}$$

The model parameters are trained using stochastic gradient descent to minimise the loss in Eq. (23). Once training is complete, the model operates over rolling windows of real-time data, forecasting both the cost components $(\widehat{\mathcal{J}}_3^t, \widehat{\mathcal{J}}_4^t)$ and the clustering parameters as $\widehat{\boldsymbol{\Theta}}_{\text{cluster}}^t = [\widehat{w}_{\text{elec}},\ \widehat{w}_{\text{sync}},\ \widehat{w}_{\text{self}},\ \widehat{\tau}_{\text{merging}},\ \widehat{\rho}_{\text{penalty}}]$.

To ensure safe operation under uncertainty, a fallback mechanism is employed. The predicted configuration is only accepted if (i) the variance of the clustering weights remains below a threshold τ_{var}, and (ii) the entropy of the latent posterior remains below τ_{entropy}. If either condition is violated, a conservative fallback $\mathbf{\Theta}^t_{\text{fallback}}$ is used:

$$\mathbf{\Theta}^t_{\text{used}} = \begin{cases} \widehat{\mathbf{\Theta}}^t_{\text{cluster}}, & \text{if both uncertainty checks are satisfied,} \\ \mathbf{\Theta}^t_{\text{fallback}}, & \text{otherwise.} \end{cases} \tag{25}$$

The fallback configuration $\mathbf{\Theta}^t_{\text{fallback}}$ reuses the most recent parameter set that successfully passed the uncertainty checks. Finally, $\mathbf{\Theta}^t_{\text{used}}$ is sent to *Step 1-1* for adaptive spectral partitioning. Also, the predicted clustering weights $(\widehat{w}_{\text{elec}}, \widehat{w}_{\text{sync}}, \widehat{w}_{\text{self}})$ and the cost estimates $(\widehat{\mathcal{J}}^t_3, \widehat{\mathcal{J}}^t_4)$ are sent to *Step 3-2*, where they help test the system's resilience to simulated cyberattacks. For simplicity, static parameters (e.g., merging threshold and merging penalty) are excluded since they do not fluctuate as much in the short term.

2) Step 3-2: Cyber-physical resilience testing and adaptive reconfiguration: The Bayesian LSTM-VAE in *Step 3-1* forecasts operational and mobility-related EV/ESS costs $(\mathcal{J}_3, \mathcal{J}_4)$ and informs clustering decisions under dynamic conditions. However, cyber-physical threats often involve rare (or unknown) attack patterns that are not captured in historical data. To address this, *Step 3-2* uses a cWGAN-GP to generate diverse and physically plausible cyberattack scenarios (e.g., FDI, DoS, and topology tampering) to assess system robustness under adversarial stress. These scenarios are used to evaluate the impact on the cyber-resilience objective $\mathcal{J}_6$ and trigger fallback reconfiguration, if necessary.

The cWGAN-GP comprises a generator $\mathcal{G}_{\text{cyber}}(\cdot)$ and discriminator $\mathcal{D}_{\text{cyber}}(\cdot)$, both conditioned on the real-time operational state vector $\mathbf{c}_t \in \mathbb{R}^{d_c}$ that encapsulates the most cyber-vulnerable features of the microgrid:

$$\mathbf{c}_t = \Bigg[\underbrace{\sum_{x_1,x_2,e} \Delta x^t_{x_1x_2,e} X^t_{x_1,x_2,e}}_{\text{EV routing}}, \quad \underbrace{\widehat{w}^t_{\text{elec}},\ \widehat{w}^t_{\text{sync}},\ \widehat{w}^t_{\text{self}}}_{\text{Clustering weights (Step 3-1)}}, \\ \underbrace{\tilde{V}^{*,t}_{b_1},\ \tilde{\theta}^{*,t}_{b_1},\ \ldots,\ \tilde{V}^{*,t}_{b_{N_b}},\ \tilde{\theta}^{*,t}_{b_{N_b}}}_{2N_b \text{ consensus boundary bus values (Step 2)}},\ \mathbf{t}_t \Bigg]^\top, \tag{26}$$

where the full set of boundary buses is given by $\mathcal{S}^{\text{boundary}} = \bigcup_p \mathcal{S}^{\text{boundary}}_p$ and the total number of unique boundary nodes is $N_b = |\mathcal{S}^{\text{boundary}}|$. These boundary values are included in the input because they are especially sensitive to cyber-physical disruptions and help capture how the system might react under attack. The temporal vector $\mathbf{t}_t \in \mathbb{R}^{d_{\text{time}}}$ includes time-related information like hour-of-day or season. All features in $\mathbf{c}_t$ are standardised using historical data. The total number of input features is $d_c = 1 + 3 + 2N_b + d_{\text{time}}$.

To train cWGAN-GP, the discriminator $\mathcal{D}_{\text{cyber}}$ maximises the difference in predicted scores between real and fake intrusion vectors, while the generator $\mathcal{G}_{\text{cyber}}$ aims to fool it. A gradient penalty is used to enforce Lipschitz continuity and improve convergence. The loss functions are defined as [24]:

$$\begin{cases} \mathcal{L}_{\mathcal{D}_{\text{cyber}}} = \mathbb{E}_{\mathbf{v}_{\text{fake}}}\left[\mathcal{D}_{\text{cyber}}(\mathbf{v}_{\text{fake}}, \mathbf{c}_t)\right] - \mathbb{E}_{\mathbf{v}_{\text{real}}}\left[\mathcal{D}_{\text{cyber}}(\mathbf{v}_{\text{real}}, \mathbf{c}_t)\right] \\ \qquad\qquad + \delta_{\text{penalty}} \cdot \mathbb{E}_{\tilde{\mathbf{v}}}\underbrace{\left[\left(\left\|\nabla_{\tilde{\mathbf{v}}}\mathcal{D}_{\text{cyber}}(\tilde{\mathbf{v}}, \mathbf{c}_t)\right\|_2 - 1\right)^2\right]}_{\text{Gradient penalty } (\mathcal{L}_{\text{penalty}})}, \\ \mathcal{L}_{\mathcal{G}_{\text{cyber}}} = -\mathbb{E}_{\mathbf{z}}\left[\mathcal{D}_{\text{cyber}}(\mathcal{G}_{\text{cyber}}(\mathbf{z}, \mathbf{c}_t), \mathbf{c}_t)\right] \end{cases} \tag{27}$$

where $\mathbf{v}_{\text{fake}} = \mathcal{G}_{\text{cyber}}(\mathbf{z}, \mathbf{c}_t) \in \mathbb{R}^{d_v}$ is a synthetic intrusion vector, generated from latent noise $\mathbf{z} \sim \mathcal{N}(0, I) \in \mathbb{R}^{d_z}$ and the current conditioning vector $\mathbf{c}_t \in \mathbb{R}^{d_c}$. The interpolated vector $\tilde{\mathbf{v}}$ used for the gradient penalty is formed as $\tilde{\mathbf{v}} = \varepsilon\mathbf{v}_{\text{real}} + (1-\varepsilon)\mathbf{v}_{\text{fake}}$, where $\varepsilon \sim \mathcal{U}(0,1)$ is a scalar drawn from the continuous uniform distribution on $[0,1]$. The operator $\mathbb{E}_x[\cdot]$ denotes the mean value taken over the distribution of the random variable x.

Once trained, the generator $\mathcal{G}_{\text{cyber}}$ can produce a wide range of intrusion vectors $\mathbf{v}_{\text{fake}}$. To ensure realism, each generated vector is passed through a physical plausibility check $\mathcal{C}_{\text{phys}}(\mathbf{v}_{\text{fake}}, \mathbf{c}_t)$, which verifies whether the operational constraints in Section II are met. Only vectors with $\mathcal{C}_{\text{phys}} = 1$ are retained for downstream testing.

Verified adversarial scenarios are used to evaluate changes in the cyber-resilience cost $\mathcal{J}_6$. To avoid overreacting to transient spikes in cyber risk from random GAN outputs, an adaptive gating mechanism is introduced. Reconfiguration is triggered only if the cyber risk increment $\Delta\mathcal{J}^t_6$ exceeds a threshold τ_{cyber} for at least N_{gate} consecutive time steps:

$$\begin{aligned} \Delta\mathcal{J}^t_6 &= \left|\mathcal{J}^{t,\text{perturbed}}_6 - \mathcal{J}^{t,\text{nominal}}_6\right| \\ &= \left|w_a\left[\Delta(\text{FPR} - \text{ADR})\right] + w_d\,\Delta\hat{\mathcal{T}}_{\text{res}} + w_u\,\Delta\text{FNR}\right|, \end{aligned} \tag{28}$$

If $\Delta\mathcal{J}^t_6 > \tau_{\text{cyber}}$, the system activates an adaptive reconfiguration policy.

$$\mathbf{\Theta}^{t,\text{updated}[1:3]}_{\text{used}} \leftarrow \mathbf{\Theta}^{t,\text{cyber}[1:3]}_{\text{fallback}}, \tag{29}$$

where $\mathbf{\Theta}^{t,\text{cyber}[1:3]}_{\text{fallback}}$ is a predefined cyber-aware clustering parameter subset selected from historical worst-case scenarios to ensure safe system operation under suspected cyber threats. It must be noted that *Step 3-1* forecasts the full clustering parameters $\mathbf{\Theta}^t_{\text{cluster}}$ for use in *Step 1-1*, but *Step 3-2* overrides the first three components by issuing the fallback set $\mathbf{\Theta}^{t,\text{cyber}[1:3]}_{\text{fallback}}$, which is then forwarded to *Step 1-1* for reconfiguration.

A potential control policy vector $\boldsymbol{\pi}^t_{\text{control}} \in \mathbb{R}^{d_\pi}$ can be defined to encode real-time operational settings that influence resilience and cost:

$$\boldsymbol{\pi}^t_{\text{control}} = \Bigg[\underbrace{r^t_{\text{EV}}}_{\substack{\text{EV routing} \\ \text{limit}}}, \quad \underbrace{P^t_{\text{ch,q}}}_{\substack{\text{EV/ESS charging} \\ \text{setpoint}}}, \quad \underbrace{P^t_{\text{dis,q}}}_{\substack{\text{EV/ESS discharging} \\ \text{setpoint}}}, \ \mathbf{t}_t \Bigg]^\top, \tag{30}$$

where r^t_{EV} controls the allowable EV routing intensity. If adversarial stress testing reveals elevated cyber risk, *Step 3-2* triggers a fallback control policy $\boldsymbol{\pi}^t_{\text{fallback}}$ that constrains these parameters to conservative values selected from a predefined library based on historical worst-case scenarios.

IV. Verification and Case Studies

This study employs the IEEE 33-bus distribution system from the MATPOWER dataset repository [25]. To model EV charging demand in the 33-bus network, static charging stations are assigned to Buses 9, 17, and 32, with baseline charging loads of 30 kW, 50 kW, and 60 kW, respectively.

Dynamic depot-level EV load profiles are superimposed using the NREL commercial fleet charging dataset [26], mapped to Bus 9 for peak demand stress, Bus 17 for moderate flexibility, and Bus 32 for low-power overnight charging. To simulate mobility-aware EV storage, the VISTA travel dataset [27] is used to generate 100 individual EV agent behaviours, mapped randomly across buses 2–33 and scheduled dynamically based on availability windows. Complementing the mobile EV agents, three fixed battery ESS are installed at Buses 7, 19, and 29 with rated capacities of 30 kWh, 40 kWh, and 35 kWh, respectively (with peak hours 17:00–21:00, up to 200 kW power capability). To evaluate the impact of these components on system stability, Fig. 2a illustrates voltage magnitudes over time at several critical buses. Three photovoltaic (PV) generators are also integrated at Buses 5, 11, and 24. To evaluate the scalability of the proposed framework, simulations are extended to the IEEE 123-bus distribution system. Six static EV charging stations are installed at Buses 15, 30, 45, 65, 85, and 100, each assigned depot-level load profiles from the NREL commercial fleet dataset. In addition, 300 mobile EV agents are synthesised using the VISTA travel behaviour dataset and randomly mapped across Buses 2–123. Five fixed ESS units, with capacities ranging from 45 to 70 kWh, are also deployed at Buses 12, 38, 52, 76, and 91. In addition, four PV generators are integrated at Buses 25, 50, 70, and 90. As illustrated in Fig. 2b, the IEEE 123-bus network exhibits significantly greater voltage variability than the 33-bus case, with some nodes (e.g., Bus 65) experiencing drops as low as 0.60 p.u. and standard deviations up to 0.057 p.u, emphasising the need for scalable coordination in systems with high EV, ESS, and PV penetration.

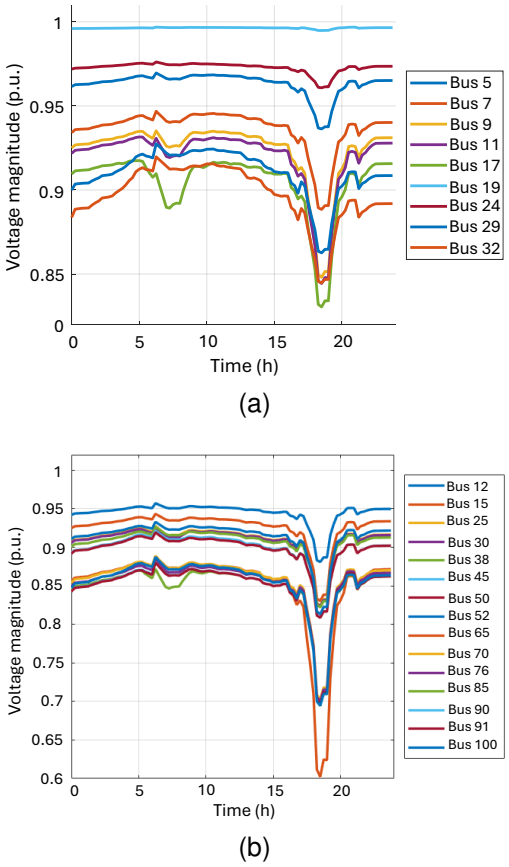


Fig. 2. Comparison of temporal voltage profiles at selected critical buses with EVs, ESS, and PVs. (a) Voltages for IEEE 33-bus. (b) Voltages for IEEE 123-bus.

1) SIAHC performance: As shown in Fig. 3a, *Step1-2* generates six initial clusters via AHC, based on a composite distance metric combining electrical proximity (w_{elec}=0.2630), dynamic synchronisation (w_{sync}=0.1121), and local self-sufficiency (w_{self}=0.6218). A conservative merging threshold (τ_{merging}=0.0598) and moderate separation penalty (ρ_{penalty}=0.2) are applied to preserve structural integrity and intra-cluster coherence. In *Step1-3*, spectral refinement is applied to each cluster. If a cluster's algebraic connectivity $\lambda_2^{(C)}$ falls below the threshold $\lambda_{\min}$= 0.07 and the cluster size $|C| \geq \tau_c$=4, it is partitioned along the Fiedler vector $\vec{v}_2^{(C)}$. As illustrated in Fig. 3b and summarised in Table II, two of the six initial clusters (2 and 4) are refined, resulting in a total of eight validated clusters. The final solution demonstrates strong clustering quality with ECI=0.8406, ICS=0.9445, and SSI=0.7089, confirming that *Step 1-3* effectively improves internal connectivity.

K-means clustering is applied to the same composite distance matrix to serve as a benchmark against SIAHC. Since K-means assumes Euclidean space, classical multidimensional scaling is used to project the non-Euclidean distances into two dimensions. To ensure a fair comparison, the number of clusters is fixed to match SIAHC's final count ($|C| = 8$). As shown in Fig. 3c, K-means fragments topologically coherent areas such as buses 14-17, due to its limited topological and structural awareness. K-means achieves higher inter-cluster separation (ICS=0.9587), but it underperforms on electrical cohesion (ECI= 0.7224) and self-sufficiency (SSI=0.5862). In terms of execution time, K-means runs significantly faster (nearly 10× faster), but it sacrifices physical interpretability.

To evaluate scalability, SIAHC is applied to the IEEE 123-bus system. As shown in Fig. 4a, *Step1-2* produces 21 initial clusters based on a composite metric (w_{elec}=0.2405, w_{sync} = 0.1373, w_{self}=0.6222), using τ_{merging}=0.044 and ρ_{penalty}=0.15. Clusters larger than 10 ($|C| \geq \tau_c$=10) with poor spectral connectivity ($\lambda_2 < 0.04$) are further refined via Fiedler vector partitioning (see Table II). This process results in 25 final clusters (Fig. 4b), with improved validation metrics: ECI=0.8871, ICS=0.9263, and SSI=0.7924. K-means is also applied using the same composite matrix, with $|C| = 25$. As depicted in Fig. 4c, K-means captures certain local groupings but causes fragmentation in topologically complex regions. Despite a slightly higher ICS = 0.9412, it shows lower cohesion (ECI = 0.7528) and self-sufficiency (SSI = 0.6121). Although K-means achieves slightly higher separation, SIAHC provides better cohesion and self-sufficiency, resulting in more meaningful and system-aware clusters.

2) ADMM performance: Fig. 5a illustrates the time-series trends of the three weighted cost components (i.e., power loss $\mathcal{J}_1$, voltage deviation $\mathcal{J}_2$, and reconfiguration $\mathcal{J}_5$), over a 24-hour horizon. All three costs rise and fall in tandem, peaking around 18:00, which shows how network congestion

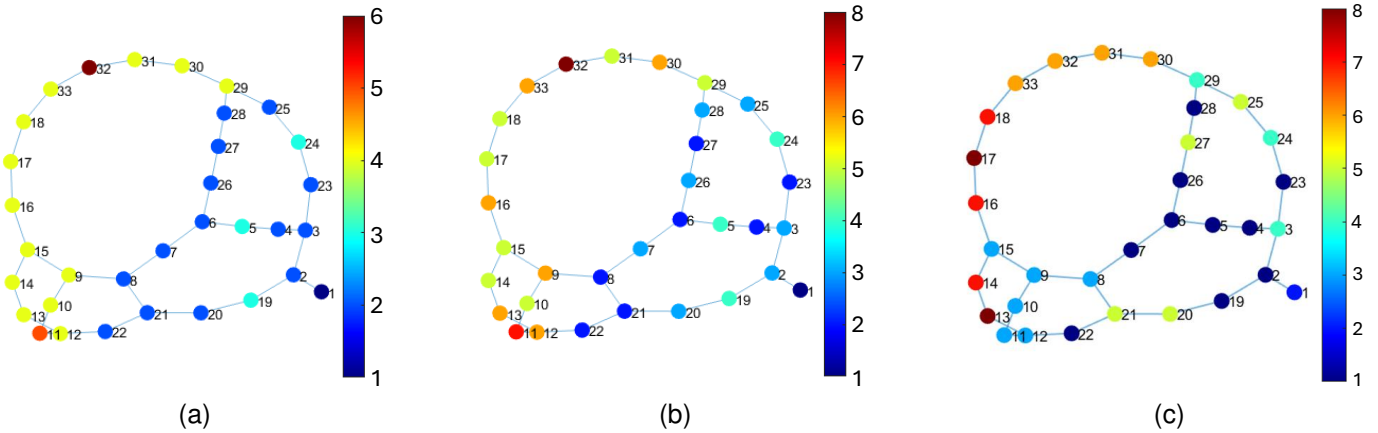


Fig. 3. Clustering results on the IEEE-33 bus system. (a) Initial clusters $\mathcal{K}$ obtained from AHC. (b) Refined clusters $\mathcal{K}_{\text{refined}}$ after applying spectral refinement. (c) Clusters obtained from K-means clustering in embedded space. (cluster colours represent group membership; colour indices may differ between subplots.)

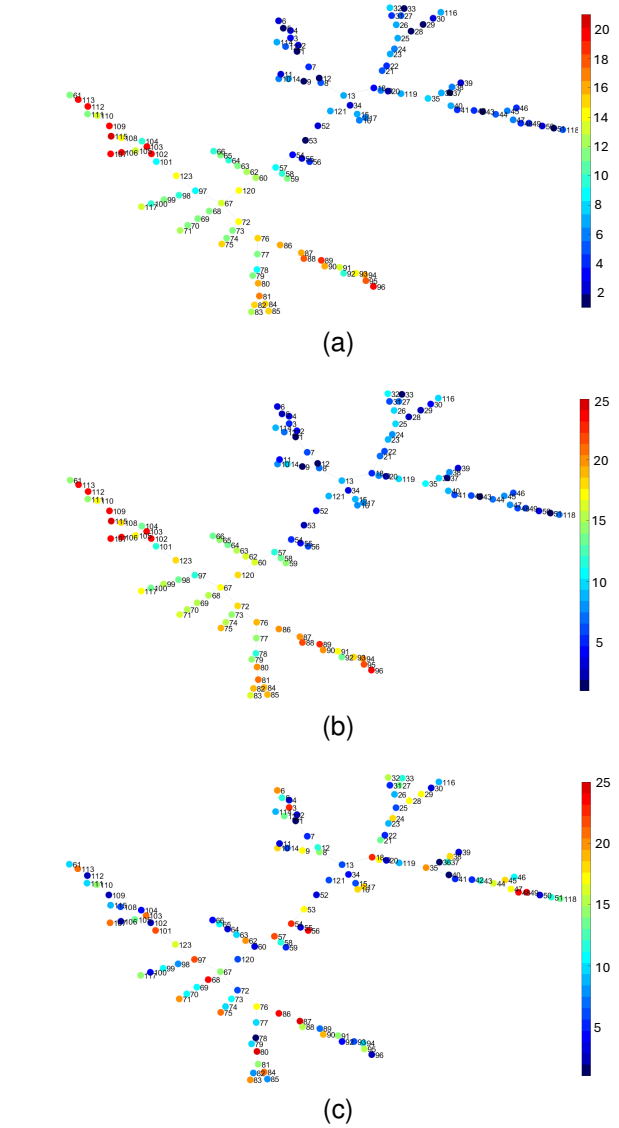


Fig. 4. Comparison of clustering results on the IEEE-123 bus system. (a) Initial clusters $\mathcal{K}$ obtained from AHC. (b) Refined clusters $\mathcal{K}_{\text{refined}}$ after spectral refinement. (c) Clusters from K-means clustering in embedded space.

TABLE II
CLUSTER REFINEMENT RESULTS
(REFINEMENT THRESHOLDS $\tau_c = 4$, $\lambda_{\min} = 0.07$ FOR IEEE-33 AND $\tau_c = 10$, $\lambda_{\min} = 0.04$ FOR IEEE-123)

System	Init	Size	λ_2	Refined?	New ID(s)	New size(s)	% Preserved
IEEE-33	1	1	NaN	No	1	1	100
	2	14	1.50×10^{-4}	Yes	2, 3	7, 7	50, 50
	3	3	3.70×10^{-14}	No	4	3	100
	4	13	6.41×10^{-7}	Yes	5, 6	7, 6	53.85, 46.15
	5	1	NaN	No	7	1	100
	6	1	NaN	No	8	1	100
IEEE-123	1	1	NaN	No	1	1	100
	2	11	1.32×10^{-4}	Yes	2, 3	5, 6	45.5, 54.5
	3	11	9.21×10^{-5}	Yes	4, 5	6, 5	54.5, 45.5
	4	3	2.88×10^{-3}	No	6	3	100
	5	9	7.02×10^{-6}	No	7	9	100
	6	6	1.43×10^{-6}	No	8	6	100
	7	8	4.11×10^{-2}	No	9	8	100
	8	12	3.54×10^{-7}	Yes	10, 11	6, 6	50, 50
	9	2	1.87×10^{-5}	No	12	2	100
	10	4	3.12×10^{-4}	No	13	4	100
	11	7	8.25×10^{-12}	No	14	7	100
	12	13	5.03×10^{-5}	Yes	15, 16	7, 6	53.85, 46.15
	13	4	2.61×10^{-2}	No	17	4	100
	14	5	3.40×10^{-6}	No	18	5	100
	15	5	7.80×10^{-7}	No	19	5	100
	16	5	6.21×10^{-5}	No	20	5	100
	17	5	9.16×10^{-4}	No	21	5	100
	18	1	NaN	No	22	1	100
	19	2	4.87×10^{-6}	No	23	2	100
	20	1	NaN	No	24	1	100
	21	8	2.78×10^{-8}	No	25	8	100

affects system-wide performance. The dominance of $\mathcal{J}_5$ is expected, given the configured weights and penalty settings ($\Gamma = [10, 10, 10]$, $w = [10, 8, 12]$, and $\alpha_\theta = 0.8$), which strongly penalise mismatches across the three reconfigurable lines ((3,8), (12,14), and (20,25)). Fig. 5b shows the cost incurred by each cluster before and after ADMM coordination. Without coordination, some clusters, particularly clusters 5 and 6, exhibit higher relative costs due to their size (or local load conditions). Once ADMM is applied, all clusters benefit from coordinated optimisation, with cost reductions most visible during peak hours. Fig. 5c compares cost transitions across four coordination stages: (i) no clustering, (ii) unrefined clustering using AHC (6 clusters), (iii) refined clustering via spectral refinement (8 clusters), and (iv) coordinated optimisation via ADMM. The comparison confirms that clustering alone yields minimal improvement, refinement enhances cluster cohesion, but it is the final ADMM-based coordination that

drives substantial cost reduction.

To address the impact of using different ADMM variants, Fig. 6a compares the performance of the conventional ADMM (fixed penalty) with other ADMM variants, all applied after the same SIAHC-based partitioning. As shown, the variable penalty ADMM used in this work (red line) achieves the lowest overall cost across most of the 24-hour horizon. Other ADMM variants also perform well during specific intervals. The convexified ADMM from [2] (pink line), which was designed for PV coordination, performs strongly between 11:00-14:00 h when PV generation is active (left inset). The smooth-consensus ADMM from [16] (blue line), which was designed for ESS coordination, shows a slight advantage during peak hours ($\approx$ 18:00-20:00 h), when storage dispatch is most active (right inset).

The ADMM-based approach in this paper demonstrates reliable and fast convergence across all evaluated scenarios. As shown in Fig. 6b, the dual residual decreases rapidly, with the variable penalty ADMM (red curve) reaching the predefined threshold ($\tau_{\text{dual}} = 10^{-3}$) by iteration 28, faster than all other methods. With simplified load conditions and fewer active constraints, the convergence is even faster and the primal residual decreased from 2.1×10^{-1} to below 10^{-4} within 8 iterations, with the dual residual meeting its threshold by iteration 14. In comparison, smooth-consensus ADMM (from [16]) converges moderately fast, while both fixed penalty ADMM and convexified ADMM (from [2]) exhibit slower convergence and more residual oscillations. The inset plot (iterations 15-45) depicts these differences more clearly, with the proposed approach showing the smoothest and steepest residual drop. These results show the advantage of adjusting the penalty parameter during the process. It starts at $\gamma^0_{\text{penalty}} = 10.0$ and gradually drops to about 0.15 by the time the algorithm converges. In addition, the use of convex relaxation for switching states and warm-started variable initialisation in this paper helps maintain stability and ensures smooth decay of residuals. Fig. 6c offers a broader estimated comparison of the four ADMM variants across eight key criteria. The variable penalty ADMM shows consistently strong performance, especially in convergence speed, residual smoothness, and final objective value. Fixed penalty ADMM stands out for its simplicity and low communication overhead, making it appealing for lightweight implementations. Smooth-consensus ADMM performs well in terms of residual smoothness and communication efficiency. The convexified ADMM outperforms others in terms of final objective value and memory usage.

The proposed ADMM-based coordination method is designed to scale efficiently with network size. As the number of clusters increases, convergence iterations and execution time grow but remain manageable due to localised, neighbour-to-neighbour communication patterns. In the IEEE-33 bus case, the algorithm converged in 28 iterations with a total execution time of 32.3 seconds, demonstrating fast and stable performance. This is owing to adaptive penalty tuning, convex relaxation of discrete control variables, and warm-started initialisation. For the more complex IEEE-123 system, which contains over three times as many clusters, convergence is achieved in 116 iterations with an execution time of 203.6 seconds.

3) Clustering parameter learning via Bayesian LSTM-VAE: The Bayesian LSTM-VAE model predicts cost components related to stationary energy storage and EV mobility ($\widehat{\mathcal{J}}_3^t, \widehat{\mathcal{J}}_4^t$), while simultaneously outputting an optimal set of clustering parameters ($\widehat{\mathbf{\Theta}}^t_{\text{used}}$), which guide the downstream process. As shown in Fig. 7, the model demonstrates strong predictive capability for key clustering weights ($\widehat{w}_{\text{elect}},\ \widehat{w}_{\text{sync}},\ \widehat{w}_{\text{self}}$) that govern affinity matrix construction. This confirms the model's ability to approximate clustering behaviour in a data-driven and interpretable manner.

The prediction accuracy for cost components $\mathcal{J}_3$ and $\mathcal{J}_4$ is assessed via root mean squared error (RMSE) and normalised mean absolute error (nMAE) over the full test horizon, and the results are presented in Table III. RMSE indicates strong accuracy, while higher nMAE reflects the influence of small true values amplifying normalised errors.

TABLE III
PREDICTION ACCURACY FOR ESS AND EV SCHEDULING COSTS

Metric	$\mathcal{J}_3$ (Storage)	$\mathcal{J}_4$ (EV scheduling)
RMSE	0.9725	1.0286
nMAE	1.0203	1.0985

4) Cyber-physical resilience via cWGAN-GP: To evaluate the resilience of the proposed GAN-aware control framework, extensive simulations are conducted on the IEEE-33 and IEEE-123 distribution systems under varying levels of EV penetration and adversarial attack intensity. Cyber threats are generated using a trained cWGAN-GP model, which produces stealthy and coordinated perturbations to voltage and power setpoints at randomly selected buses, mimicking FDI attacks. To construct the resilience heatmaps shown in Fig. 8, test scenarios are defined across combinations of EV penetration levels and attack intensities (i.e., numbers of buses compromised). For each scenario, a series of Monte Carlo simulations is performed as follows: the cWGAN-GP generates perturbations, the partitioned control scheme responds, and the system's voltage profiles are recorded. A constraint violation is registered if the voltage magnitude at any bus exceeds the safety range ($V_{\min} \leq V_b^t \leq V_{\max}$). The fraction of control clusters with at least one violation is computed and averaged across 10 trials to ensure statistical robustness. The resulting heatmaps illustrate the system's vulnerability surface, where darker regions indicate a higher proportion of clusters violating voltage constraints. The dashed boundary indicates the fallback threshold, empirically defined as the point where more than 5% of clusters exhibit voltage violations, triggering reconfiguration. Within this boundary, the system remains stable without intervention, while crossing it signals the onset of systemic stress. Results show that in the IEEE-33 system, constraint violations rise significantly beyond 7 attacked buses and 22% EV penetration (see Fig. 8a). In contrast, the IEEE-123 system tolerates up to 22 simultaneous attacks and 40% EV penetration with minimal disruption (see Fig. 8b). This demonstrates the advantages of larger system scale, greater control granularity, and a more complex topology in absorbing adversarial disturbances.

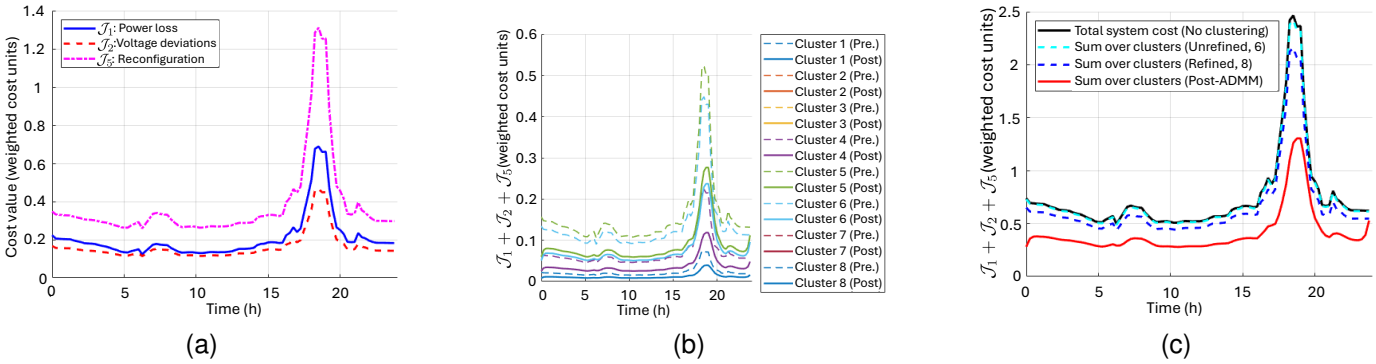


Fig. 5. The impact of ADMM-based optimisation on cost trends across coordination stages and time. (a) Weighted cost components over time. (b) Per-cluster cost comparison before and after ADMM coordination. (c) Cost transitions for: no clustering, AHC clustering, refined clustering, and ADMM-based optimisation.

To determine an appropriate threshold for triggering fallback reconfiguration (τ_{cyber}), a sensitivity analysis is conducted (Fig. 9). As τ_{cyber} increases, fallback activations become less frequent (Fig. 9a), but the deviation $\Delta\mathcal{J}6^{\text{norm}}$ (Fig. 9b) grows post-attack. The average cost $\bar{\mathcal{J}}6$ (Fig. 9c) shows a minimum near $\tau_{\text{cyber}} \approx 0.28$, balancing system stability and responsiveness. Detection metrics in Fig. 9d–9e show expected behaviour: increasing the threshold reduces the FPR but raises the FNR. These opposing effects are summarised by a customised resilience score $\mathcal{R}_{\text{cyber}}$, which is defined as:

$$\mathcal{R}_{\text{cyber}} = \frac{\text{ADR} - \text{FPR}}{\hat{\mathcal{T}}_{\text{res}} + \Delta\mathcal{J}_6^{\text{norm}} + \varepsilon_{\text{cyber}}}, \tag{31}$$

where $\varepsilon_{\text{cyber}} > 0$ avoids division by zero. This score integrates detection accuracy (captured by high ADR and low FPR in the numerator), control responsiveness (penalised by response delay), and post-fallback stability (penalised by the residual cost deviation) into a single metric. A higher $\mathcal{R}_{\text{cyber}}$ indicates a more resilient system that not only detects attacks reliably but also responds swiftly and stabilises effectively. As can be seen from Fig. 9f, it apeaks near $\tau_{\text{cyber}} \approx 0.28$, confirming it as the most balanced and effective threshold.

Fig. 10 shows the normalised deviation in the cybersecurity cost $\Delta\mathcal{J}_6^{\text{norm}}$ under adversarial perturbations. The dashed blue line represents the raw signal, while the solid black curve is the smoothed version used for detection. The fallback policy $\boldsymbol{\Theta}_{\text{fallback}}^{t,\text{cyber}[1:3]}$ is triggered whenever $\Delta\mathcal{J}_6^{\text{norm}}$ exceeds the optimal threshold $\tau_{\text{cyber}} = 0.28$ (green dashed line). Two attack intervals (around $t = 35-45$ and $t = 55-65$) are detected, prompting fallback responses. Blue markers indicate invalid operating points observed during these disturbances. After each fallback activation, the system quickly stabilises, with the post-fallback mean (magenta line) converging near 0.035.

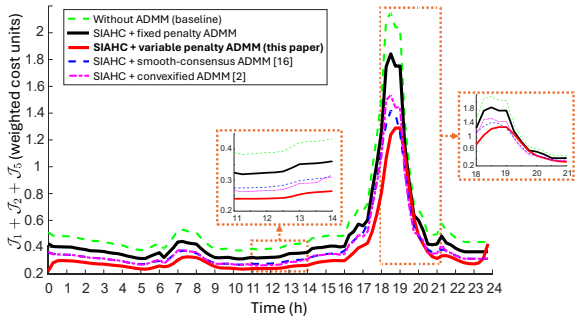
Fig. 11 illustrates the cyber anomaly detection performance of the cWGAN-GP across three major attack categories: FDI, DoS, and topology manipulation. These categories are defined as: (i) FDI involves the injection of false voltage or power setpoints, (ii) DoS represents communication delays or blocking between control agents, and (iii) topology tampering (or manipulation) refers to falsified reports of line-switching states or network connectivity. The ADR remains consistently high, ranging from 88% (DoS) to 95% (topology), demonstrating the model's strong ability to identify malicious behaviour. FPR and FNR remain low, especially under topology manipulation (FPR = 0.05, FNR = 0.03), demonstrating precise detections. The results confirm the discriminator's robustness and the framework's capacity for a timely, accurate fallback response.

A quantitative summary of cyber-resilience indicators is presented in Table IV. The average response delay following attack detection $\hat{\mathcal{T}}_{\text{res}}$ is 2.1 time steps, and 5 reconfiguration events are triggered in total, confirming the responsiveness of the fallback mechanism. The low post-fallback deviation of ≈ 0.035 further demonstrates the stabilising impact of the control policy.

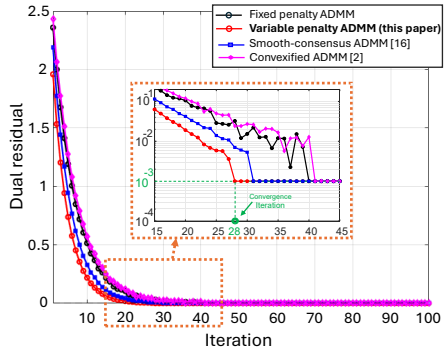
TABLE IV
RESILIENCE AND CYBERSECURITY RESPONSE METRICS.

Metric	Value	Interpretation
ADR	0.94	High detection quality
FPR	0.06	Low alert noise
FNR	0.03	Few missed intrusions
$\hat{\mathcal{T}}_{\text{res}}$	2.1 steps	Fast fallback action
Triggered reconfigs	5	System adaptivity
Post-fallback mean	0.035	Stabilised cost

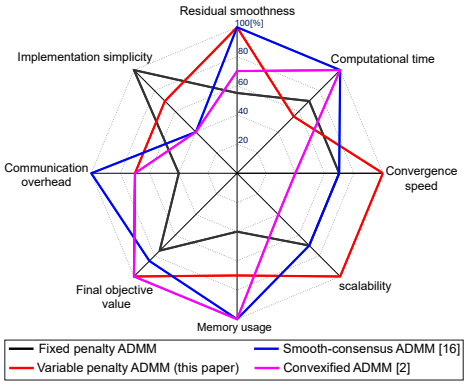
To further evaluate the effectiveness of the proposed GAN-aware resilience framework, a comparative analysis is conducted against two representative hybrid methods from Table I: the cloud-edge framework in [17] and the dynamic partitioning strategy for shared storage in [16]. These approaches reflect partial alignment with the proposed framework through their use of data-driven detection and distributed optimisation, respectively. In [17], deep learning is integrated with partitioning rules to support reconfiguration under abnormal load behaviour. However, it does not explicitly model coordinated cyber threats or maintain structural interpretability, thereby limiting its applicability in adversarial scenarios. The method in [16] focuses on optimal storage allocation using multi-objective optimisation and decision-making, supporting adaptability under renewable variability but assuming nominal conditions without any cyber-physical threat modelling or recovery logic. In contrast, the proposed framework integrates cWGAN-based adversarial modelling, LSTM-based forecast-



(a)



(b)



(c)

Fig. 6. Performance comparisons between different ADMM variants. (a) Weighted cost over 24 hours. (b) Convergence trends over 100 iterations. (c) Multi-metric evaluation of ADMM variants

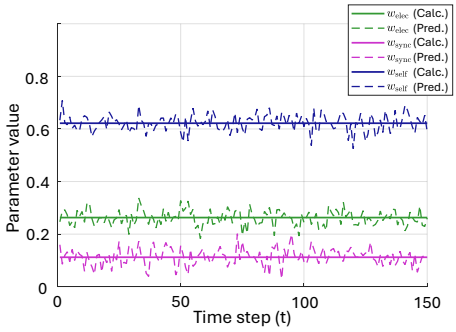
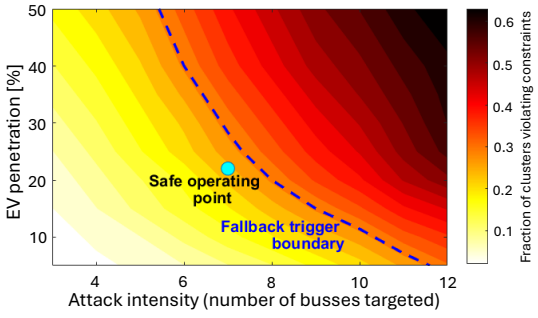
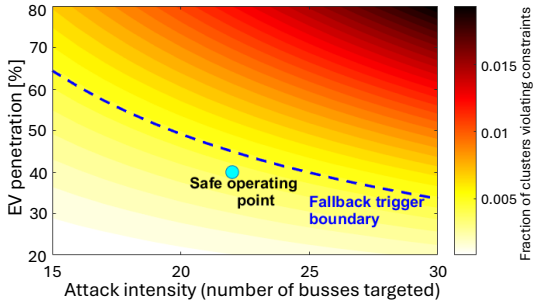


Fig. 7. Clustering weights $\widehat{w}_{\text{elec}}$, $\widehat{w}_{\text{sync}}$, and $\widehat{w}_{\text{self}}$ learnt from the Bayesian LSTM-VAE versus extracted values for SIAHC.



(a)



(b)

Fig. 8. Resilience heatmaps showing the fraction of clusters violating voltage constraints under varying EV penetration levels and GAN-induced cyberattacks. (a) IEEE 33-bus system results. (b) IEEE 123-bus system results. (darker regions indicate higher vulnerability.)

ing, and ADMM-driven partitioned control. This combination enables proactive adaptation and resilience assessment under evolving threat and load dynamics. Its performance is quantified using the resilience score $\mathcal{R}_{\text{cyber}}$ defined in Eq. 31. Values for [17] and [16] are estimated based on their design assumptions and reported behaviour. As shown in Fig. 12, the proposed framework achieves a higher score across varying EV penetrations. It preserves high resilience up to the dropp-off threshold of 35% EV penetration before a controlled decline, due to built-in anticipation and fallback logic. This nonlinear trend shows that the framework can handle moderate stress using forecasting and reconfiguration, until it reaches the limit of its control capacity. The framework in [17] achieves a moderate resilience score of approximately 0.12, which reflects its support for adaptive reconfiguration but without fallback or post-attack recovery. The framework in [16], which does not implement any detection or reconfiguration mechanisms, maintains a flat and low score throughout. As EV penetration increases, [17] exhibits a steady linear decline, while [16] shows no significant change due to its fixed control design.

V. Limitations and Future Work

The proposed framework demonstrates strong empirical performance, but several important limitations remain open for future work. Although the ADMM-based coordination enables parallel solution of local subproblems, it does not overcome the non-convexity inherent in AC power flow and operational constraints. Classical ADMM does not offer formal convergence guarantees in such settings. To address this, convex

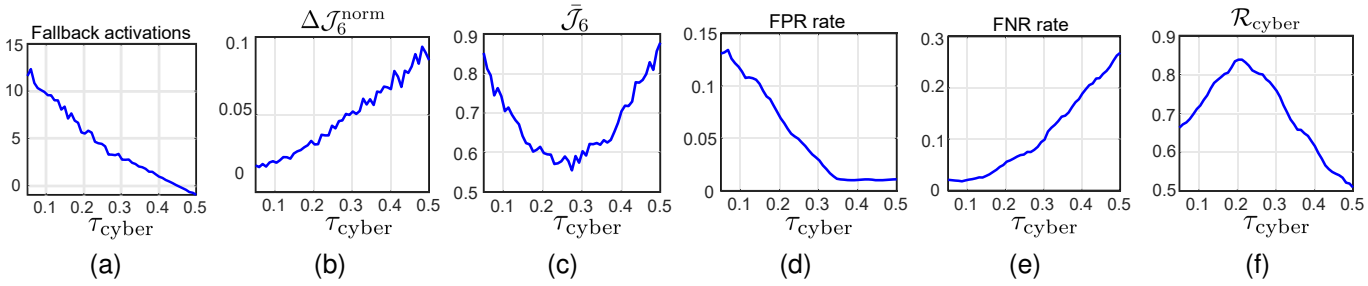


Fig. 9. Sensitivity analysis of cyber-physical resilience metrics versus the cybersecurity threshold τ_{cyber}. (a) Fallback activations. (b) Post-fallback deviation. (c) Average resilience cost. (d) False positive rate. (e) False negative rate. (f) Resilience score.

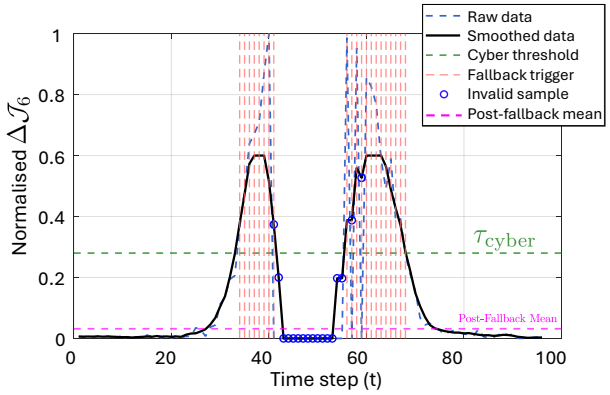


Fig. 10. Normalised deviation in cybersecurity cost $\Delta\mathcal{J}_6^{\text{norm}}$ over time.

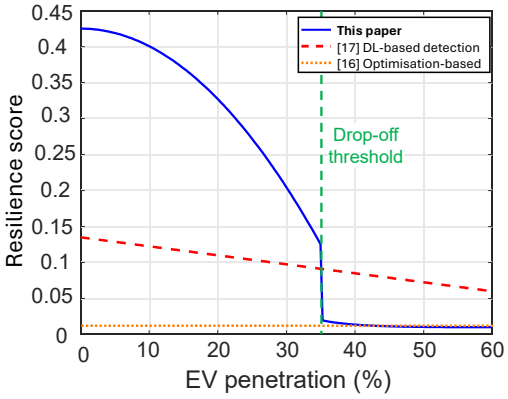


Fig. 12. Comparisons: resilience score $\mathcal{R}_{\text{cyber}}$ vs EV penetration.

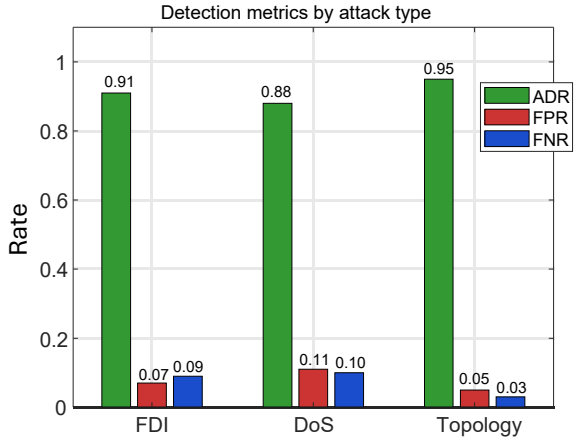


Fig. 11. Cyber anomaly detection performance across different attack types using cWGAN-GP.

relaxation is applied to binary switching variables, and an adaptive penalty update is used to stabilise residuals. However, convergence remains heuristic, and future work should explore more robust non-convex optimisation techniques such as proximal-ADMM or semidefinite programming to improve theoretical guarantees. Also, the resilience layer primarily addresses cyber-physical threats (e.g., FDI attacks) but does not currently consider other types of disturbances, such as hardware faults or communication failures. Besides, resilience metrics like detection accuracy, fallback delay, and cost deviation are only assessed in post-simulation analysis, rather than being incorporated directly into real-time optimisation. Integrating such metrics into the control layer could support more proactive and adaptive system responses. Additionally, the current EV model treats vehicles as mobile storage units with fixed availability profiles. This abstraction enables tractable partitioning and control but overlooks user behaviour and mobility patterns, which could significantly affect operational decisions. More realistic, behaviour-driven EV models will be considered in future work. Finally, although the proposed architecture achieves effective coordination under a centralised model, where clustering and fallback actions are handled by the distribution system operator, it relies on system-wide information and does not employ real-time inter-cluster communication. Each partition operates semi-independently, which limits decentralised adaptability. A promising future direction is to investigate peer-to-peer or decentralised architectures in which clusters can autonomously share local state and negotiate reconfiguration decisions, improving scalability, responsiveness, and resilience under practical communication constraints.

VI. Conclusions

This paper presents a structured hybrid partitioning and coordination framework to improve the operability and cyber-physical resilience of distribution networks under dynamic conditions. The proposed pipeline integrates spectral graph clustering, distributed ADMM-based optimisation, Bayesian forecasting, and generative adversarial stress testing to enable real-time control, reconfiguration, and cybersecurity awareness

at scale. Simulation results on the IEEE 33-bus and 123-bus systems validate the framework's effectiveness. The SIAHC clustering method with spectral refinement yields modular, self-sufficient partitions that maintain topological integrity and outperform conventional techniques such as K-means in electrical cohesion and interpretability. The variable-penalty ADMM scheme supports efficient, stable optimisation, significantly reducing power losses, voltage deviations, and reconfiguration costs, particularly during peak conditions, while maintaining scalability. The Bayesian LSTM-VAE model enables cost-relevant clustering parameters forecasting with high accuracy (RMSE < 1.03), supporting adaptive partitioning. Meanwhile, the cWGAN-GP-based cyber-resilience layer identifies vulnerability thresholds and simulates stealth attacks. When triggered, the fallback mechanism restores system stability rapidly (within 2.1 time steps on average), with post-attack cost deviation remaining below 0.035. Resilience heatmaps and sensitivity analyses confirm the framework's robustness across varying cyberattack intensities and EV penetration levels. Compared to other hybrid methods, it consistently achieves superior composite resilience scores.

ACKNOWLEDGMENTS

The work is supported by the U.K. Leverhulme Trust grant RPG-2023-107.

REFERENCES

[1] S. Jena and N. P. Padhy, "Cyber-Secure Global Energy Equalization in DC Microgrid Clusters Under Data Manipulation Attacks," *IEEE Trans. Ind. Appl.*, vol. 59, no. 5, pp. 5488-5505, 2023.

[2] Y. Chai, L. Guo, C. Wang, Z. Zhao, X. Du and J. Pan, "Network Partition and Voltage Coordination Control for Distribution Networks With High Penetration of Distributed PV Units," *IEEE Trans. Power Syst.*, vol. 33, no. 3, pp. 3396-3407, May 2018.

[3] S. K. Panda and B. Subudhi, "A Review on Robust and Adaptive Control Schemes for Microgrid," *J. Mod. Power Syst. Clean Energy*, vol. 11, no. 4, pp. 1027-1040, July 2023.

[4] F. Mohammadi et al., "Robust Control Strategies for Microgrids: A Review," *IEEE Syst. J.*, vol. 16, no. 2, pp. 2401-2412, 2022.

[5] Y. Han, K. Zhang, H. Li, E. A. A. Coelho and J. M. Guerrero, "MAS-Based Distributed Coordinated Control and Optimization in Microgrid and Microgrid Clusters: A Comprehensive Overview," *EEE Trans. Power Electron.*, vol. 33, no. 8, pp. 6488-6508, 2018.

[6] D. M. Manias et al., "Trends in Smart Grid Cyber-Physical Security: Components, Threats, and Solutions," *IEEE Access*, vol. 12, pp. 161329-161356, 2024.

[7] J. Ye et al., "A Review of Cyber–Physical Security for Photovoltaic Systems," *IEEE J. Emerg. Sel. Topics Power Electron.*, vol. 10, no. 4, pp. 4879-4901, Aug. 2022.

[8] S. M. Miraftabzadeh, C. G. Colombo, M. Longo and F. Foiadelli, "K-Means and Alternative Clustering Methods in Modern Power Systems," *IEEE Access*, vol. 11, pp. 119596-119633, 2023.

[9] M. Massaoudi, M. Ez Eddin, A. Ghrayeb, H. Abu-Rub and S. S. Refaat, "Advancing Coherent Power Grid Partitioning: A Review Embracing Machine and Deep Learning," *IEEE Open Access J. Power Energy*, vol. 12, pp. 59-75, 2025.

[10] M. Mao, Z. Wu, D. Xu, J. Xu and Q. Hu, "Community-Detection-Based Approaches for Distribution Network Partition," *CSEE J. Power Energy Syst.*, vol. 10, no. 5, pp. 1965-1976, Sept. 2024.

[11] A. R. D. Fazio, C. Risi, M. Russo and M. D. Santis, "Coordinated Optimization for Zone-Based Voltage Control in Distribution Grids," *IEEE Trans. Ind. Appl.*, vol. 58, no. 1, pp. 173-184, 2022.

[12] G. Pierrou, H. Lai, G. Hug and X. Wang, "A Decentralized Wide-Area Voltage Control Scheme for Coordinated Secondary Voltage Regulation Using PMUs," *IEEE Trans. Power Syst.*, vol. 39, no. 6, pp. 7153-7165, Nov. 2024.

[13] Y. Bansal, R. Sodhi, S. Chakrabarti and A. Sharma, "A Novel Two-Stage Partitioning Based Reconfiguration Method for Active Distribution Networks," *IEEE Trans. Power Del.*, vol. 38, no. 6, pp. 4004-4016, Dec. 2023.

[14] R. J. Sánchez-García et al., "Hierarchical Spectral Clustering of Power Grids," *IEEE Trans. Power Syst.*, vol. 29, no. 5, pp. 2229-2237, Sept. 2014.

[15] J. Ge, Z. Wu, J. Xu, and Q. Hu, "A two-stage flow-based partition framework for unbalanced distribution networks," *CSEE J. Power Energy Syst.*, pp. 1–11, 2023.

[16] J. Li, Z. Fang, Q. Wang, M. Zhang, Y. Li and W. Zhang, "Optimal Operation with Dynamic Partitioning Strategy for Centralized Shared Energy Storage Station with Integration of Large-scale Renewable Energy," *J. Mod. Power Syst. Clean Energy*, vol. 12, no. 2, pp. 359-370, March 2024.

[17] R. Wang et al., "A Cloud-Edge Intelligence-Based Optimization Method for Distribution Network Partitioning and Operation Considering Simulation Inaccuracy," *IEEE Trans. Power Syst.*, pp. 1-13, Jan. 2025.

[18] Y. Wang, L. Lebovitz, K. Zheng and Y. Zhou, "Consensus Clustering for Bi-objective Power Network Partition," *CSEE J. Power Energy Syst.*, vol. 8, no. 4, pp. 973-982, July 2022.

[19] R. J. Sánchez-García et al., "Hierarchical Spectral Clustering of Power Grids," *IEEE Trans. Power Syst.*, vol. 29, no. 5, pp. 2229-2237, Sept. 2014.

[20] X. Xu, F. Xue, S. Lu, H. Zhu, L. Jiang and B. Han, "Structural and Hierarchical Partitioning of Virtual Microgrids in Power Distribution Network," *IEEE Syst. J.*, vol. 13, no. 1, pp. 823-832, 2019.

[21] J. Wu, X. Chen, S. Badakhshan, J. Zhang and P. Wang, "Spectral Graph Clustering for Intentional Islanding Operations in Resilient Hybrid Energy Systems," *IEEE Trans. Industr. Inform.*, vol. 19, no. 4, pp. 5956-5964, April 2023.

[22] M. Ez Eddin, M. Massaoudi, H. Abu-Rub, M. Shadmand and M. Abdallah, "Optimum Partition of Power Networks Using Singular Value Decomposition and Affinity Propagation," *IEEE Trans. Power Syst.*, vol. 39, no. 5, pp. 6359-6371, Sept. 2024.

[23] X. Zhang et al., "Reactive Voltage Partitioning Method for the Power Grid With Comprehensive Consideration of Wind Power Fluctuation and Uncertainty," *IEEE Access*, vol. 8, pp. 124514-124525, 2020.

[24] S. Li, H. Xiong and Y. Chen, "DiffCharge: Generating EV Charging Scenarios via a Denoising Diffusion Model," *IEEE Trans. Smart Grid*, vol. 15, no. 4, pp. 3936-3949, July 2024.

[25] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "MATPOWER: Steady-State Operations, Planning, and Analysis Tools for Power Systems Research and Education," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 12–19, Feb. 2011.

[26] National Renewable Energy Laboratory. (2021). Fleet Depot Charging Profiles Dataset. NREL Data Catalogue., [Online]. Available: https://data.nrel.gov/submissions/162.

[27] Y. Wu, S. M. Aziz, M. H. Haque, "Travel Datasets for Analysing The Electric Vehicle Charging Demand in a University Campus," *Data in Brief*, Vol. 54, p. 110335, 2024.