**PhD Oral Examination Proposal**

**CAPSM: Cognitive Adaptive Power System Management**

**A Brain-Inspired Dual-Process Architecture for Modern Grid Control**

Mahmoud Kiasari

Department of Electrical and Computer Engineering

Dalhousie University, Halifax, Nova Scotia

Supervisor: Dr. Aly

Co-Supervisor: Dr. Gu & Dr. Hammad

November 2025

**Abstract -** Modern power systems comprise an increasing number of cyber-physical and renewable resources, which creates a fundamental controversy between the need for rapid responses from the protection and optimality planning levels in grid control. Existing methodologies, such as Model Predictive Control (MPC), Optimal Power Flow (OPF), Deep Reinforcement Learning (DRL), Safe Reinforcement Learning (Safe RL), Physics-informed Neural Networks (PINNs), and Multi-agent Reinforcement Learning (MARL), have their own ideas for one part of this problem, but none provide a coherent physics-consistent architecture that would be able to provide a mixture of millisecond fault reflexes, cost-optimal operation, and cyber-resilient coordination at scale. This proposal presents a Cognitive-Adaptive Power System Management (CAPSM) framework inspired by dual-process theories of human cognition. CAPSM consists of (i) a System-1 layer, which is based on fast CNN-LSTM controllers and PINN-constrained policies providing a reflexive protection-like reaction; (ii) a System-2 layer, which is based on quantum-inspired Reinforcement Learning (QIRL) and hierarchical multi-agent coordination providing slower, globally optimal decision-making; and (iii) a metacognitive arbitration layer, which can switch dynamically between the two systems and combine them according to uncertainty, novelty, and risk. Building on the previous work of this applicant on EV/V2G control, Flexible AC Transmission System (FACTS)-based power quality enhancement, and AI-enhanced grid management, CAPSM will be developed and validated on the occurrence of the most foundational standards in power system studies as study cases in the case of the system based on 9-, 39-, 118-, and 300-bus systems with high renewable penetration, FACTS devices, and V2G fleets. The research will assess the ability of CAPSM to (a) provide a sub 5 ms response of System 1 and sub 50 ms planning of System 2; (b) have a very low constraint violation rate (goal <0.2%); (c) improve economic performance and power quality compared to MPC and DRL baselines; and (d) maintain its resilience under cyber-attacks and communication delays. The anticipated end result is a brain-inspired, physics-aware, and metacognitively adaptive control paradigm that is a direct response to the research gaps highlighted in the literature and can be used as a basis for next-generation smart grid intelligence

Table of Contents

[1.Introduction and Research Motivation 5](#_Toc215603204)

[1.1 Emerging Challenges in Modern Power Systems 5](#_Toc215603205)

[1.2 The Need for Unified, Intelligent and Physics Incorporating Control Paradigm 6](#_Toc215603206)

[2. Literature Review and Limitations of Existing Approaches 6](#_Toc215603207)

[2.1 Model Predictive Control and OPF Based Supervisory Control 7](#_Toc215603208)

[2.2 Grid Control using Deep Reinforcement Learning 7](#_Toc215603209)

[2.3 Safe Reinforcement Learning in Power System 8](#_Toc215603210)

[2.4 Physics-Informed Neural Network and Physics-Constrained Learning 8](#_Toc215603211)

[2.5 Multi-Agent Reinforcement Learning for Energy Networks 9](#_Toc215603212)

[2.6 Integration of Internet of Things and Artificial Intelligence 9](#_Toc215603213)

[2.7 EV/V2G Control and Grid-Integrated Optimization 10](#_Toc215603214)

[2.8 Structural Gaps in the Literature 10](#_Toc215603215)

[3. Research gaps, Problem Statement, and Hypothesis 11](#_Toc215603216)

[3.1 Research Gaps 11](#_Toc215603217)

[3.2 Problem Statement 13](#_Toc215603218)

[3.3 Central Hypothesis 13](#_Toc215603219)

[3.4 Research Objectives 14](#_Toc215603220)

[Expected Outcome 16](#_Toc215603221)

[4. Proposed Framework: Cognitive Adaptive Power system Management (CAPSM) 18](#_Toc215603222)

[4.1 Conceptual Overview 18](#_Toc215603223)

[4.2 Integration Depth vs. Component Novelty 19](#_Toc215603224)

[4.3 System-1: Reflexive Intelligence 20](#_Toc215603225)

[4.3.1 Architecture and Operation 20](#_Toc215603226)

[4.3.2 Functional Role 21](#_Toc215603227)

[4.4 System-2: Deliberative Optimization 21](#_Toc215603228)

[4.4.1 Quantum-Inspired Reinforcement Learning (QIRL) 21](#_Toc215603229)

[4.4.2 Control Functions 22](#_Toc215603230)

[4.5 Metacognitive Arbitration Layer 23](#_Toc215603231)

[4.5.1 Decision Logic 23](#_Toc215603232)

[4.5.2 Adaptive Coordination 24](#_Toc215603233)

[4.5.3 Robustness under Uncertainty 24](#_Toc215603234)

[4.6 Physics-Informed and Hierarchical Design 24](#_Toc215603235)

[4.6.1 Physics Informed Neural Networks (PINNs) 24](#_Toc215603236)

[4.6.2 Hierarchical Multi-agent Coordination (H-MARL) 25](#_Toc215603237)

[4.6.3 Multi way Fusion and Uncertainty Quantification 26](#_Toc215603238)

[4.7 Existing Foundations and Preliminary Work 26](#_Toc215603239)

[5. Research Methodology and Implementation Plan 27](#_Toc215603240)

[6. Preliminary Simulation Results and Validation 29](#_Toc215603241)

[6.1 System-1 Reflexive Fault Detection Performance 29](#_Toc215603242)

[6.2 System-2 QIRL Optimization vs Classical DQN 31](#_Toc215603243)

[6.3 Physics-Informed Safety and Constraint Compliance (PINNs) 31](#_Toc215603244)

[6.4 HMARL Scalability with System Size and Agent Count 32](#_Toc215603245)

[6.5 Real-Time Response Characteristics 33](#_Toc215603246)

[6.6 Computational Resources and Asynchronous Operation 33](#_Toc215603247)

[6.7 Cyber-Physical Resilience 34](#_Toc215603248)

[7. Future Work and Timeline 34](#_Toc215603249)

[8. Conclusions 36](#_Toc215603250)

[References: 37](#_Toc215603251)

[Appendix: 38](#_Toc215603252)

# 1.Introduction and Research Motivation

The electrical power grid is experiencing an unprecedented transformation into a cyber-physical ecosystem that is rich in renewable energy; thus, the fundamental principles that govern power grid operation, protection, and control are being challenged. Renewable energy sources, mainly wind and solar, currently account for more than 30% of the worldwide electric energy generation capacity. The International Energy Agency (IEA) expects this to increase to more than 50% by 2035. Simultaneously, the traffic electrification trend is gaining momentum, and the Electrical Vehicle (EV) fleet will exceed 300 million units by 2030, fundamentally changing the demand pattern and both grid and demand characteristics. Recently, the development of distributed energy resources (DERs), including energy storage tie-in points, microgrids, and prosumer-based generation, has resulted in substantial bidirectional power flows and decentralized decision-making, which are unprecedented in conventional power-grid systems.

Consequently, the grid has transformed over time from a centrally controlled, unidirectional energy delivery system to a heterogeneous, data-intensive, and highly dynamic cyber-physical system (CPS). However, this transition also presents some fundamental problems in ensuring stability, security, and real-time control under conditions of extreme variability and uncertainty.

## 1.1 Emerging Challenges in Modern Power Systems

The increasing complexity of renewable networks, often found in long-distance networks, has led to a series of multilayered technical and operational challenges that challenge the traditional control paradigm. Intermittency and Uncertainty Because the energy generated from solar and wind is intermittent by design, the output can vary by 50-80% within minutes because of weather variability [4]. Such stochastic behavior presents the problem of instability in frequency and deviation in voltage and requires rapid and continuous balancing actions. Traditional deterministic or short-term predictive models, such as those based on numerical weather forecasts, fail to maintain their accuracy under extreme conditions or massive ramp developments, which often cause curtailment, load shedding, or expensive reserve activation [5]– [6].

* **Heterogeneous Complexity:** While legacy systems feature robust networks of a few hundred controllable devices, modern grids comprise several thousand smart electronics, including FACTS controllers, smart inverters, battery storage, and EV charging. Each device interacts with the others through nonlinear feedback mechanisms, and communication delays exponentially increase the decision space, making centralized optimization (intractable in principle).
* **Speed-optimality Trade-off:** Protective control in the event of a fault or cascading failure requires the system to be able to react in milliseconds, for example, less than 5 ms to avoid collapsing the voltage or tripping the lines. However, optimizing the global economy and operations requires large-scale optimization computations that typically take hundreds of milliseconds to several seconds. This basic trade-off between speed and optimality is one of the most prolonged and unanswered dilemmas in power system control [9].
* **Cybersecurity and Communication Resilience:** The growing interconnection between operational technology (OT) and information technology (IT) systems exposes the grid to cyberattacks, data injection, and denial-of-service attacks, thereby compromising situational awareness and control reliability. Conventional control architectures typically assume reliable data input channels and are thus susceptible to malicious or incorrect data inputs.
* **Physical and Operational Constraints:** All control actions must comply with physical laws (Kirchhoff’s laws), voltage magnitude limits, thermal ratings and stability margins. Violations of these constraints may cause equipment degradation, protection maloperation, and ultimately, cascading blackouts. The key limitation of most existing AI-based control systems is their inability to ensure constraint satisfaction in the presence of uncertainty, noise, and nonlinear dynamics.

## 1.2 The Need for Unified, Intelligent and Physics Incorporating Control Paradigm

The coexistence of rapid dynamic events, large-scale optimization requirements, and physical constraints is insufficient in existing hierarchical control architectures. Traditional supervisory and protection layers operate independently, often in conflicting modes with divergent timescales. Centralized Model Predictive Control (MPC) makes mathematically optimal decisions but is computationally intensive (usually 500-2000 mas) and is prone to errors in modelling, especially under high renewable penetration and topological changes. Deep Reinforcement Learning (DRL) techniques, such as Deep Q-Network (DQN) or Proximal Policy Optimization (PPO), can deal with faster responses (tens of milliseconds) at the cost of slow convergence, low sampling efficiency, and videos without explanation of enforcing any constraints. Meanwhile, rule-based protection systems are limited in their reaction speed (sub-millisecond) and adaptability to novel fault situations and cyber-physical disturbances.

These shortcomings highlight the urgent need for a unified, cognitively inspired control architecture that can coordinate fast protection, slower optimization, and safety mechanisms across grids. Instead of treating relays, FACTS devices, and dispatch algorithms as isolated modules, the framework should jointly consider dynamic stability, economic efficiency, and cyber-physical security, while explicitly arbitrating between fast reflexive actions and slower deliberative decisions.

# 2. Literature Review and Limitations of Existing Approaches

Modern power systems operate at the intersection of renewable intermittency, cyber-physical coupling, and distributed intelligence, under which traditional protection and control frameworks struggle to deliver real-time stability, safety, and optimal performance. Although substantial work has been conducted on model-based optimization, reinforcement learning, physics-informed modelling, and multi-agent control, the majority of existing approaches remain **isolated**, **single-capability**, or **non-scalable**. This section synthesizes the dominant research directions from 2018 to 2025 and identifies their structural weaknesses, forming the foundation for the research gaps addressed by the CAPSM framework, as follows:

## 2.1 Model Predictive Control and OPF Based Supervisory Control

MPC and Optimal Power Flow (OPF) are the basic models for the optimal control and dispatching of power systems. MPC predicts the system evolution over a finite horizon and optimizes the control inputs given the prescribed constraints, whereas OPF is responsible for optimizing the steady-state power injections and flows in the network. In the field of power electronics, Karamanakos et al. surveyed different MPC variants for converters and drives, which can provide better tracking performance and explicit treatment of the current and voltage constraints of medium-voltage power electronic systems[1]. Within the context of microgrids, Shahzad et al. provided a good summary of the main MPC strategies applied to islanded and grid-connected microgrids, which are useful for voltage control, frequency support, and economic dispatching [2]. Recently, some MPC formulations for networked microgrids that can be scaled to the network size have been proposed. For example, Jiang et al. used decomposition and distributed optimization techniques to extend the applicability of MPC to large networks [3]. Notwithstanding these advances, MPC and OPF have intrinsic structural limitations when examining a renewable-rich cyber-physical grid [3]. First, their computational latency is typically incompatible with protection-level response times, although some advanced solvers and model-reduction techniques have been developed. The solution times for realistic 39- or 118-bus systems have been reported to range from hundreds of milliseconds to several seconds. Fault-induced voltage dips and converter overcurrent’s require fault mitigation within a few milliseconds. Consequently, the MPC is inherently suited to supervisory control and not to sub cycle reflexes. Second, MPC performance is closely linked to the accuracy of the models employed, and any changes in the topology, device outages, or fast renewable ramp events may render the linearization’s or reduced-order models underlying the MPC invalid, leading to deteriorated tracking or loss of stability. Third, centralized formulations of the OPF and MPC do not scale well as the network size and complexity continue to grow, and although decomposition schemes help ease this problem, they generally compromise global optimality or increase implementation complexity. Finally, both MPC and OPF assume the integrity of measurements and communications; they have no built-in mechanisms to detect or counter false data injection or coordinated cyber-attacks, making them vulnerable when placed in the position of a primary decision-maker in an adversarial setting [4], [5]. These limitations imply that MPC/OPF must continue to be employed as invaluable supervisory tools; however, they cannot, by themselves, provide the unity of real-time cognitive control envisioned by CAPSM.

## 2.2 Grid Control using Deep Reinforcement Learning

The era of DRL has spawned many studies aimed at replacing or complementing model-based control with data-based policies. In the field of power systems, DRL has been applied for voltage control, load shedding, and emergency stabilization. Hossain et al. proposed a model-based DRL framework (i.e., MB-PARS) for emergency voltage control in large-scale grids using the case study of the IEEE 300 bus and demonstrated significant improvements in the sample efficiency and training time compared to purely model-free DRL approaches [6]. Other investigations consider the DRL for inverter dispatch, demand response, and frequency regulation, making use of the fact that RL agents can learn by interacting with nonlinear simulators without explicitly deriving the models [7], [8]. However, DRL faces critical difficulties when considered as a contender for real-time security and grid supervision. The most important challenge is in terms of sample efficiency and cost of training; even the more sample-efficient versions using the model require many simulated episodes to converge, especially when the policy is of high dimension on a realistic test system [6]. Second, most DRL formulations enforce safety and constraint satisfaction using soft penalties in the reward function, which may mean that during training, the voltage, current, or line loading constraint is often violated, making it infeasible to allow DRL algorithms to be directly deployed into operational systems. Third, DRL policies are usually trained on fixed operating regimes and are not robust to distributional shifts, such as changing generation portfolios, new grid configurations, and cyber-physical disturbances. Policies that perform well in one set of scenarios may fail catastrophically if there are unseen contingencies. Finally, although DRL controllers may have satisfactory empirical performance, they are often opaque; thus, they offer no explicit guarantees on stability, constraint satisfaction, or interpretability, thus breaking the confidence of operators for mission-critical applications. These limitations highlight the consideration that DRL as a sole means cannot be used as the "brain" of a safety-critical cognitive control architecture.

## 2.3 Safe Reinforcement Learning in Power System

To address the safety limitations of standard DRL, recent studies on safe reinforcement learning have introduced mechanisms such as constrained Markov decision processes, shielded policies, Lyapunov-based critics, and model-predictive safety filters. These approaches can maintain trajectories within predefined safety sets or soften constraint violations through penalties. However, most safe RL schemes for power systems remain proof-of-concept: they are typically validated on low-dimensional test cases, assume simplified dynamics, or rely on hand-tuned safety sets that do not scale to realistic transmission networks. Moreover, they rarely integrate domain knowledge, such as protection settings, voltage and reactive power limits, or explicit multi-timescale reasoning. This leaves a gap between the theoretically safe RL and practical physics-aware controllers for large power grids [9].

## 2.4 Physics-Informed Neural Network and Physics-Constrained Learning

Physics-informed neural networks (PINNs) were proposed by Raissi, Perdikaris, and Karniadakis as a method to incorporate governing equations, usually partial differential equations (PDEs) and ordinary differential equations (ODEs), into the loss function of a neural network to allow the network to approximate solutions that respect physical laws. Since their introduction, PINNs and their extensions have been used in multiple domains of fluid dynamics, material science, and more recently, in power systems for power flow approximation and dynamic state estimation. Recent surveys have identified PINNs as a core tool in a scientific machine learning approach that helps rewrite the liabilities of making explicit connections between empirical data and mechanistic models. Within the power systems field (and other fields), PINN-like approaches have been used for approximating the alternating-current power flow and transient dynamics, with respect to the enforcement of Kirchhoff's laws and device-level equations in the training procedure [10]. Nevertheless, most implementations are still offline, serving primarily to accelerate simulations, refine state estimations, or perform sensitivity analyses rather than to produce real-time control actions. The direct integration of PINNs into a real-time controller raises several new problems because we must pay the computational overheads of computing the physics residuals and use stiff dynamics with discrete control policies. Furthermore, extant studies focus on single-timescale problems to the point that the coordination of millisecond time-scale protective reflexes with slower second time-scale supervisory decisions has been overlooked. CAPSM proposes to go beyond the use of PINNs as auxiliary approximations and incorporate physics into the decision-making process, where constraint residuals and physical feasibility are integral parts of the cognitive evaluation process in both fast and slow layers.

## 2.5 Multi-Agent Reinforcement Learning for Energy Networks

With the development of power systems to reduce their level of centralization, MARL has become an interesting paradigm for coordinating distributed energy resources, microgrids, and smart buildings. A recent survey on MARL for energy networks was conducted by Keren et al. [11]. This emphasizes the potential of MARL as an alternative to centralized schemes for distributed decision-making in markets, microgrids, and transactive energy systems (ES). Concomitantly, many studies have been conducted on MARL for specific energy management problems. For example, Wilk et al. developed a strategy-based local optimization of the energy resources of a smart city energy management, where the buildings were implemented by multiple agents [12]. Other studies have used MARL for residential energy management, multi-microgrid cooperation, and peer-to-peer transactive energy trading, and have shown better scalability and privacy compared to centralized control. However, MARL for energy networks introduces new challenges. Training stability decreases as the number of agents increases and the environment becomes more non-stationary. Most MARLs implemented for energy systems typically operate at comparatively slow timescales, focusing on scheduling and dispatch instead of supporting transient stability and protection. In addition, most MARL formulations do not explicitly incorporate physics, but instead have their simulators or simple models, and physical constraints are used as punishment or to check the requirements. Consequently, actions derived from agents may be physically infeasible or unsafe unless post-processing or other safety filters are applied. Finally, MARL research almost never systematically deals with cyber-physical threats or adversarial actors. While MARL provides a scalable framework for distributed intelligence, it does not, in and of itself, provide the structure of physics consistency, multi-timescale consistency, cognitive processing, and mindful capabilities that CAPSM understates itself as having the possibility of providing it.

## 2.6 Integration of Internet of Things and Artificial Intelligence

The unification of communication networks, Internet of Things (IoT) devices, and intelligent metering has significantly expanded the attack surface of power systems. A comprehensive review by Ghiasi et al. addressed cyber-attacks and defense mechanisms on smart grids, which cover false data injection, denial of service, and coordinated cyber-physical threats in smart grids from theoretical and practical perspectives. Ghiasi et al.. They pointed out that as control becomes increasingly automated and data-driven, sophisticated attacks against the measurement and control channels may cause stealthy taxi failures. Ghadi et al. built on this perspective by proposing security risk models that leverage big data and artificial intelligence to quantify and mitigate cyber risks in smart grid operations [4]. Most studies on this topic focus on detection and risk assessment rather than on cognitively adaptive control. Intrusion detection systems, anomaly detectors, and risk-scoring mechanisms detect such suspicious behaviors; however, the controllers downstream, regardless of whether they are MPC, DRL, or rule-based schemes, do not usually change their way of reasoning or level of confidence in response to such signals. In effect, although modern grids often have monitoring "eyes" but they lack the integrative "brain," brain that can change the strategy under attack. CAPSM explicitly models cyber-physical security signals into cognitive states, allowing for metacognitive arbitration between immediate protective responses, caution supervision, or recovery strategies when cyber uncertainty is high.

## 2.7 EV/V2G Control and Grid-Integrated Optimization

Electric vehicles (EVs) and vehicle-to-grid (V2G) systems pose both risks and opportunities for the operation of electric grids. Large-scale charging can lead to peak demand aggravation and cause local congestion; however, EVs can also act as distributed storage, providing frequency support, congestion, and renewable firming. MPC remains the dominant methodology for scheduling and controlling the charging and V2G active participation of EVs. Hermans et al. present an implementation study of MPC based control for vehicle charging stations in a grid-connected microgrid to demonstrate the feasibility of predictive charging under real-world complexities in the transformers, feeders and tariffs [13]. Other studies have examined robust and stochastic MPC for EV fleets with arrival time, initial state-of-charge, and other driving pattern uncertainties [13]. Although these strategies are effective in managing EV charging (at the station or fleet level), they are typically considered subsystems and do not operate as integral parts of a unified control structure, that is, a grid-wide cognitive controller. EV controllers typically assume fixed network conditions and grid codes and rarely communicate with higher-level protection or optimization mechanisms in real time. Moreover, most EV/V2G MPC implementations operate on a timescale of minutes and cannot address fast transient support on faults or fast renewable fluctuations. CAPSM anticipates EVs not simply as ideally scheduled storage units, but as part of a cognitive ecosystem, in which they are likely to participate continuously in slow planning and occasionally reflexively on call.

## 2.8 Structural Gaps in the Literature

A coherent picture emerges across these strands of literature. Model-based methods, such as MPC and OPF, offer mathematically elegant supervisory control but are too slow, brittle, and centralized for real-time protection and high uncertainty. Although DRL introduces adaptability, it sacrifices safety, interpretability, and robustness to distribution shifts. Safe RL mitigates some of these issues but remains constrained to limited timescales and relatively simple constraint structures. PINNs and physics-informed learning provide a powerful mechanism for embedding physical laws but are rarely integrated into decision-making loops, particularly in multi-timescale settings. MARL offers a path to scalable decentralization but is largely physics-agnostic and sensitive to nonstationarity. Cybersecurity research has produced sophisticated detection and risk assessment tools; however, the insights from these systems have not yet been embedded into the control logic in a cognitively meaningful way. Although EV/V2G control is advancing rapidly, it is still treated as a separate, siloed subsystem rather than an active participant in a unified control architecture.

*Table 1 Research Scope and Boundaries of the CAPSM PhD Proposal*

|  |  |  |
| --- | --- | --- |
| **Category** | **In Scope** | **Out of Scope** |
| **Dual-Process Control Architecture** | Conceptual design and mathematical formulation of a dual-process (System-1 / System-2) control framework for power systems. | Detailed neurocognitive modelling beyond what is needed for control design. |
| **Quantum-Inspired Reinforcement Learning** | Development and evaluation of QIRL algorithms for grid operation (dispatch, V2G scheduling, FACTS coordination). | Quantum hardware implementation; full quantum computing theory. |
| **Physics-Informed Neural Networks (PINNs)** | Design of PINN-based safety layers and constraint-aware policies embedded in grid controllers (voltages, currents, line limits, SoC limits). | General-purpose PINN theory unrelated to power systems. |
| **Hierarchical Multi-Agent RL (HMARL)** | Design and testing of hierarchical multi-agent coordination for FACTS devices, distributed generation, and EV/V2G fleets. | Generic MARL benchmarks (e.g., games) unrelated to power systems. |
| **Test Systems & Domains** | IEEE 9-, 39-, 118-, and 300-bus systems with high renewables, FACTS, and V2G participation; representative cyber-attack and communication-delay scenarios. | Full national-scale market models; detailed tariff design and regulatory analysis. |
| **Hardware / Real-Time Platforms** | Software-based co-simulation (MATLAB/Simulink + Python) and consideration of real-time constraints; conceptual path to HIL implementation. | Full hardware prototyping or industrial HIL deployment. |
| **Power Quality & Cyber-Physical Resilience** | Indices such as voltage profiles, frequency deviations, THD, line loading, SoC trajectories; resilience under disturbances, cyber-attacks, and communication delays. | Detailed electromagnetic transient (EMT) modelling of every component. |
| **Prior Work Integration** | Reuse and extension of the candidate’s existing work on EV/V2G control, FACTS-based power quality enhancement, and AI-enhanced grid forecasting and management. | Treating previous papers as independent, unrelated case studies. |

# 3. Research gaps, Problem Statement, and Hypothesis

The literature reviewed in Chapter 2 shows significant advances in MPC/OPF-based supervision, DRL and safe RL, physics-informed learning, multi-agent control, cybersecurity, and EV/V2G coordination. However, these contributions remain largely isolated from each other. No existing framework unifies fast protection, multi-timescale optimization, physics-consistent learning, multi-agent scalability and cognitive adaptability under cyber-physical uncertainty. This section formalizes the key research gaps, central problem statement, and hypothesis that guide the proposed CAPSM framework.

## 3.1 Research Gaps

**Gap 1 – Separation of fast protection and slower optimization across timescales**

Contemporary architectures of power system controls divide protection and optimization into two levels: capabilities incorporated into local reception networks, such as relay or local inverter protection I, act at millisecond levels; capabilities embodied in model predictive control, as well as optimal power flow and safe reinforcement learning soreness, operate within overview time frames of tens of milliseconds to several minutes. The coordination among these strata is generally controlled by static heuristics or offline-tuned logic, which rules out a principled mechanism that ensures the correspondence between fast reflexive actions and long-term optimization goals. This temporal dissociation becomes acute in scenarios involving rapid renewable power ramps, topological alterations, or cyber-physical perturbations. The resulting protective actions can subvert economic dispatch and optimization strategies, which are inherently too slow to prevent cascading failures.

**Gap 2 – Lack of physics-embedded learning for real-time decision-making**

Most existing machine learning controllers for power systems treat the grid as a black box; they learn statistical mappings from measurements to actions without explicitly enforcing Kirchhoff’s laws, network constraints, or protection limits in real time. Physics-informed neural networks (PINNs) and related physics-constrained learning methods have been applied mainly to offline tasks such as state estimation, surrogate modelling, and forecasting. To date, no framework that embeds AC power-flow equations, protection margins, and thermal limits directly inside the decision-making loop of real-time controllers has been widely deployed. This gap motivates the integration of physics-informed constraints into both the fast reflexive and slower planning layers of the CAPSM.

**Gap 3 – Absence of metacognitive adaptability in reinforcement learning controllers**

Standard reinforcement learners and even safe RLs have no explicit representation of self-assessment regarding reliability. Policies do not track confidence, novelty, or distribution shifts, which leads to the inappropriate application of nominal-condition policies in regimes that have not been previously observed, such as severe faults, cyber-attacks, or extreme renewable conditions. The existing literature does not provide a power-system-specific framework that translates the metacognitive concepts of confidence evaluation, uncertainty awareness, and mode switching into a control architecture that can arbitrate between fast-acting reflexive actions and planning.

**Gap 4 – Limited scalability and coordination in multi-agent reinforcement learning for power systems**

Research on multi-agent reinforcement learning and distributed optimization has been conducted in the domains of microgrids, smart communities, and energy markets; however, the lack of scalability becomes apparent when the number of agents and the non-stnonstationaritye environment increase. MARL training often undergoes instability and tends to oscillate or diverge. Many studies overlook hard physical constraints or penalize them as soft constraints, risking violations of voltage, current, or thermal constraints during both training and deployment. Moreover, few frameworks of MARL have provably convergent or stable behavior when instantiated on large test systems, such as the network erschutterten by the benchmark networks of the IEC 118- or 300-and-250bus net under realistic communication latency and data loss.

**Gap 5 – Fragmented treatment of cyber-physical security in control design**

Cybersecurity research related to smart grids has led to the development of advanced attack models and detection mechanisms; however, such knowledge is typically separated from the actual control logic. Anomaly detectors or intrusion detection systems detect suspicious data, but downstream controllers (either MPC, DRL or rule based) are not architected to modify their way of reasoning in light of heightened levels of cyber uncertainty. There is no unified architecture in which cyber-risk indicators, physical consistency checks, and control decisions are evaluated jointly in a cognitive style, thereby avoiding the system changing to a safe mode of operation when data integrity is compromised.

Collectively, these gaps highlight the need for a unifying, cognition-inspired control system built on physically consistent reflexes and opponent[s]-coordination optimized planning with metacognitive arbitration and scalable multi-architecture/agent coordination.

## 3.2 Problem Statement

Given the deficiencies identified in the prior discussion, the central research problem of the present study is as follows:

*How can a unified dual-process control architecture be developed for modern power systems that simultaneously provides both millisecond-scale protective reflexes and multi-timescale optimal decision-making while embedding the physics of power systems, including enabling power-system metacognitive adaptability to uncertainty and scaling to extensive, renewable-rich, cyber-physical networks?*

This problem explicitly addresses the age-old debate between speed and optimality in grid control, the need for physics-consistent learning, and the need for scalable, cyber-resilient coordination in future smart grids.

## 3.3 Central Hypothesis

The principal hypothesis of this research is that a dual-process metacognitively guided, physics-informed control architecture (with a rapid, reflexive layer (System-1) and a slower deliberative layer (System-2)) can achieve responsiveness and safety, and enable scalability in renewable dominant power systems simultaneously by directly integrating physical constraints within learning algorithms and dynamically arbitrating between reflexive and deliberative control modes of action under metacognitive uncertainty.

* **Dual-process cognition**: The separation of fast, intuitive reactions and slower, analytical reasoning in human cognition can be mapped to power system control, where System-1 handles sub-millisecond protective actions and System-2 handles global optimization and planning.
* **Metacognitive arbitration:** An executive layer that monitors confidence, novelty, and urgency can decide when to rely on System-1, when to invoke System-2, and how to blend their influence, thereby stabilizing the closed-loop dynamics in the face of disturbances and cyber-physical uncertainty.
* **Physics-informed learning:** Embedding Kirchhoff's laws, voltage and current limits, and stability constraints into the learning and control processes ensures that decisions remain physically feasible and interpretable, even under data noise and model uncertainty.

## 3.4 Research Objectives

To test the hypothesis, the research design was organized around five interdependent objectives that cover different technical layers of the proposed framework.

**Objective 1 - Develop a Cognitive Control model of Power Systems**

A formal mathematical model will be developed to translate dual-process cognition mappings and power system operating applications. The grid dynamics will be modelled as an MDP with state-occupying action and reward formulations for the goal of short-term protection and long-term optimization. Occurrences in analytical timing demarcate shifts between System‑1 (rapid control, < 5 ms) and System‑2 (optimized planning, 10-100 ms) actions. Theoretically, bounded delay and convergence are proved by Lyapunov basis stability analysis and time-domain performance measures.

Here, is the state space (e.g., bus voltages, angles, frequencies, SoC values), and is the action space (e.g., reactive power setpoints, FACTS control inputs, EV charging/discharging commands).  
denotes the transition kernel , is the reward (or negative cost) function, and is the discount factor that controls how much future performance matters relative to immediate performance.

Here, is the system state at time , is the control action applied by CAPSM, and represents exogenous disturbances such as renewable fluctuations, load changes, or faults.  
The function encapsulates the underlying grid physics (e.g., power-flow and dynamic equations) and defines how the state evolves in response to control actions and disturbances.

The MDP reward is designed to penalize constraint violations and inefficient operations.

Here, and are weighting matrices that penalize deviations in state (e.g., voltage deviations, frequency deviations, SoC imbalances) and large control actions, respectively.  
scales the penalty term , which aggregates violations of operational limits such as voltage bands, line current ratings, and SoC bounds.

CAPSM explicitly separates fast reflexive control from slower deliberative optimization.

Here, is the characteristic timescale of System–1 (fast, protective reflexes) and denotes the slower timescale of System–2 (deliberative optimization and planning).

The inequality formalizes the idea that reflexive actions must respond at protection-like speeds, while planning can operate on slower control cycles.

The stability of the closed-loop CAPSM controller is expressed by a Lyapunov function as follows:

Here, is a Lyapunov candidate function that measures the “energy” or deviation of the system from a desired equilibrium (e.g., nominal voltages and frequency).

The inequality with guarantees that decreases over time, ensuring that the combined effect of System–1 and System–2 leads to asymptotic stability.

**Objective 2 - Design Quantum Inspired Reinforcement Learning (QIRL) Algorithms**

A new QIRL will be developed to maximize the exploration/exploitation balance using superposition/interference/tunneling for quantum dynamics. Unlike conventional DRL (e.g., DQN, PPO), the QIRL agent operates on complex-valued probability amplitudes to enhance the sample efficiency and stability of convergence. Benchmark tests on systems with nine and 39 buses according to the standard revision of the IEEE will measure benefits in terms of convergence speed, energy cost reduction, and constraint satisfaction under stochastic disturbances.

**Objective 3 - System Physics Informed Neural Networks (PINNs) for Safety and Compliance to Constraints**

To ensure that the control actions are physically valid, Physics-Informed Neural Networks are embedded in both the reflexive and deliberative layers of the CAPSM. These networks represent AC power flow equations, voltage magnitude limits, and thermal constraints as differentiable penalty terms in the training loss. The goal of the approach is to achieve higher than 99% compliance with operational constraints without post hoc projection to preserve the real-time capability while still allowing for interpretability and reliability under varying renewable inputs.

**Objective 4 - Implement Multimodal Fusion and Hierarchical Multi-agent Coordination**

A multimodal data fusion engine will be developed to integrate heterogeneous data streams of information from phasor measurement units (PMU) (30-120 Hz), Supervisory Control and Data Acquisition (SCADA) (2-4 s), and renewable forecasts using probabilistic uncertainty quantification. A hierarchical HMARL structure is used to manage distributed entities such as generators, FACTS devices, EV fleets, and DER clusters. Agents communicate using graphs of consensus-based communication and offer scalability and resiliency in the face of communication delays or partial data loss.

**Objective 5 - Testing the CAPSM Framework through the IEEE Test Systems**

Comprehensive validation will be carried out via co-simulation in MATLAB/Simulink and Python-based environments. IEEE 9-, 39-, 118-, and 300-bus test systems were used to evaluate the performance of CAPSM under scenarios of renewable intermittency, cyber-attacks, and component faults. Metrics will include:

* Reaction time (should be less than 5 ms for System-1, and less than 50 ms for System-2)
* Constraint-violation rate (target<0.2%)
* Energy cost and loss savings compared to MPC and DRL baselines
* Fault detection accuracy/recovery time
* Scalability with the growth of the network size and agent number

Experimental outcomes will be statistically analyzed (e.g., paired t-tests, p < 0.01) to ensure the significance of the observed improvements.

Constraint-violation rate over a horizon :

Here, is the indicator function that equals 1 when any voltage or current constraint is violated at time , and 0 otherwise.

The metric therefore measures the fraction of time steps where the system operates outside secure limits, allowing direct comparison between CAPSM and baseline controllers.

Average response time to disturbances:

Here, is the number of fault or disturbance events, is the onset time of event , and is the time when system variables return within acceptable tolerance bands.

The metric quantifies how quickly the controller restores stability after disturbances.

Relative cost reduction compared to the baseline:

Here, is the total operating cost (or aggregated performance index) under a conventional controller, and is the corresponding cost under CAPSM.

expresses the relative improvement as a percentage, directly communicating the economic benefits to system operators and reviewers.

## Expected Outcome

The proposed research is expected to provide an explainable, cognitive-inspired control architecture that unifies protection, optimization, and learning in one framework. In particular, the CAPSM framework should:

* A dual-process model of control in the brain with formally analyzed properties of stability in which System-1 of the brain supports ballistic reflexive response over milliseconds and System-2 generates slower responses but better planning.
* QIRL-based controllers for finding the convergence of solutions with higher quality and faster convergence rates than standard DRL algorithms when solving non-convex stochastic grid optimization problems.
* Physics-informed safety mechanisms: Maintaining voltages, currents, and power flows within acceptable bounds of large disturbances and cyber-physical uncertainties.
* Hierarchical multi-agent layer to support the coordinated operation of FACTS devices, distributed generation, and EV/V2G fleets of large-scale networks.

These results will be tested against quantitative performance criteria using well-established test systems.

* Dynamic response: sub-5 milliseconds decision latency to change ( Reaktion) target System-1. sub-50 ms decision latency target System-2 in protection relevant scenarios.
* Constraint satisfaction: Target constraint violation rates below 0.2 % %i2%ll operating scenarios, including explicit control of voltage/frequency and line loading, as well as SoC limits.
* Economic and power quality benefits Source electricity for needs Measure reduction in operating cost Minimize losses Minimize power quality violations compared to the baselines of MPC/OPF and DRL
* Learning performance: The learning performance of the convergence speed, stability, and sample efficiency of QIRL and those of DQN, PPO, and related algorithms were compared.
* Scalability and resilience: Investigate performance degradation as network size and the number of agents increase, as well as robustness in the case of communication delays and cyber-attack situations.

Table 2 CAPSM Contribution Framework – Components, Gaps Addressed, and Evaluation

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **CAPSM Component / Layer** | **Role in CAPSM** | **Main Literature Gaps Addressed** | **Expected Contribution (Design Goals, not pre-claimed results)** | **Evaluation / Metrics** |
| **Dual-Process Architecture (System-1/2)** | Provide joint protection-level reflexes (System-1) and planning-level optimization (System-2). | Lack of unified protection–planning framework; weak treatment of timescales. | A practical, mathematically grounded dual-process control structure for power systems with explicit timescale separation and stability analysis. | Decision latency (ms vs tens of ms), stability margins, performance under disturbances. |
| **Metacognitive Arbitration Layer** | Monitor uncertainty/risk and switch/blend between System-1 and System-2 policies. | No metacognitive coordination in existing DRL/MPC frameworks. | Metacognitive supervisor that uses uncertainty/novelty indicators to decide when to rely on fast reflexes vs deliberative optimization. | Frequency of switches, performance under distribution shift, robustness to unforeseen events. |
| **QIRL-Based System-2 Planning** | Solve non-convex, stochastic grid optimization (dispatch, V2G, FACTS) using quantum-inspired RL. | Slow or unstable convergence of classical DRL; sensitivity to local minima. | QIRL algorithms tailored to grid problems that **aim** to converge faster and more reliably than DQN/PPO while handling stochasticity and constraints. | Convergence speed, final cost, robustness across scenarios vs DQN/PPO/AC OPF/MPC baselines. |
| **PINN-Based Safety and Constraint Layers** | Embed physical laws and operational limits directly into learning and control policies. | Limited integration of physics into learning; reliance on penalty methods. | Physics-informed safety shields that **target** high constraint satisfaction (≈99%+) and reduce unsafe actions without excessive penalty tuning. | Constraint violation rate, number of unsafe actions, recovery time after disturbances. |
| **Hierarchical Multi-Agent RL (HMARL)** | Coordinate FACTS, DG, and EV/V2G agents across network layers using RL with communication structure. | Poor scalability of centralized methods; MARL without physical constraints. | Scalable HMARL architecture that coordinates many agents under realistic communication/topology constraints, while respecting grid physics via PINNs and constraints. | Performance vs number of agents, sensitivity to communication delays/loss, power-quality indices. |
| **Cyber-Physical Monitoring & Defence** | Integrate anomaly detection and cyber-attack scenarios into the control design and evaluation. | Cybersecurity often treated separately from control; limited co-design. | Co-designed control and monitoring strategies where CAPSM’s decisions are tested under cyber-attack and communication disruption scenarios. | Resilience metrics (loss of load, voltage violations, recovery time) under attack vs baseline controllers. |
| **Integration of Prior Work** | Use existing EV/V2G, FACTS, and AI forecasting work as building blocks inside CAPSM. | Fragmented treatment of EV/FACTS/AI contributions. | A unified framework that **extends** the candidate’s previous results rather than duplicating them, demonstrating continuity and cumulative impact across multiple publications. | Reuse of validated models, incremental performance improvements over previous published schemes. |

# 4. Proposed Framework: Cognitive Adaptive Power system Management (CAPSM)

The CAPSM framework presents a new brain-inspired control structure that aims to unify protective reflexes and deliberative optimization in power system operations. CAPSM mimics human cognitive structures (more precisely, dual-process intelligence) to combine the fast, instinctive responses required for system protection with the slow, strategic thinking required to optimize the system as a whole. The architecture is united at a higher cognitive level in a cohesive control paradigm of data-driven learning, physical modelling, and cognitive arbitration.

## 4.1 Conceptual Overview

CAPSM is conceptually inspired by the dual-process theory of cognition, which distinguishes between two complementary systems of reasoning:

* **System 1 (Reflexive layer):** fast, automatic, and intuitive.
* **System - 2 (Deliberative layer):** Slow, Analytical, and Logical

Translating this psychological model into a power system control context, System 1 is designed to provide emergency responses (e.g., faults, rapid voltage dips, or frequency deviations) on the millisecond scale, whereas System 2 is for higher-order optimization at broader spatial and temporal scales. At its core, CAPSM models the entire control environment as an MDP, where the states are the grid instantaneous electrical conditions (voltages, currents, frequency, and power flows), actions are feasible control signals, and rewards are the stability and performance goals of the system. In this MDP formulation, System-1 and System-2 are complementary agents with different temporal scales of interaction and a common metacognitive supervisor. The metacognitive arbitration layer is "**executive function**" that identifies at any time which of the two control types to be engaged depending on the properties of the event (criticality, uncertainty, computational load, and latency budget). Through hierarchical integration, CAPSM ensures that emergent contingencies are reacted to quickly through rapid reflexes, while long-term efficiency and constraint compliance are upheld by deliberative optimization. Therefore, such an architecture creates a continuum of control from milliseconds to seconds, addressing decades of separation between the protection and optimization of today's power grids.

CAPSM implements two complementary policies, one for each cognitive system:

Here, represents the stochastic policy of the fast reflexive controller (System–1), mapping the current state to a distribution over actions .  
is the policy of the slower deliberative controller (System2), which may use more context and planning but reacts less frequently.

The metacognitive layer selects the active policy at each time step as follows:

Here, is a discrete indicator choosing between System–1 () and System–2 () based on confidence, novelty, and urgency metrics.

The effective CAPSM policy is therefore a mixture policy where the arbitration mechanism decides whether to prioritize speed or optimality.

## 4.2 Integration Depth vs. Component Novelty

The originality of the CAPSM framework is not based on the invention of new components but rather on the depth of integration. Its primary novelty is the integration of various advanced technologies into a common cognitive control architecture. Specifically, CAPSM offers four novel contributions: (1) a metacognitive arbitration mechanism that harmonizes reflexive (System-1) and deliberative (System-2) decision-making infrastructures on overlapping timescales; (2) embedding PINNs within both control loops to ensure real-time constraint satisfaction; (3) adapting QIRL to accelerate convergence and optimize within complex power system landscapes; and (4) presenting a unified reward infrastructure that prevents conflicts between fast protection and slow optimization objectives. Collectively, these integrations form the first cognitive dual-process architecture applied to power system control, emphasizing intelligent adaptability, physical consistency, and dynamic stability.

In contrast, CAPSM employs a number of existing technologies, e.g., CNN-LSTM models for PMU data, H-MARL-based frameworks such as PyMARL and RLlib, device models based on the IEEE standards FACTS, and existing formulations for V2G/G2V. CAPSM does not use such technologies as novel inventions but as enabling components for validation and scalability. The depth of the framework has been demonstrated with more than three peer-reviewed journal papers, formal convergence and Lyapunov stability proofs, ablation and complexity analysis, unlike other approaches, as well as statistically significant results (p < 0.01) results comparison with baseline controllers. The resulting interactions among System-1, System-2, the metacognitive layer, and the physics-informed/HMARL components are summarized in the CAPSM overview diagram shown in Figure 1.

A diagram of a company

AI-generated content may be incorrect.

Figure 1 CAPSM Framework Overview Diagram

## 4.3 System-1: Reflexive Intelligence

System-1 represents the protective reflex of the CAPSM architecture. It is responsible for ultra-fast responses (< 5 ms) to transient disturbances, ensuring voltage and frequency stability during sudden contingencies.

### 4.3.1 Architecture and Operation

System-1 employs a Convolutional Long Short-Term Memory (CNN-LSTM) network trained on high-frequency sensor streams, including PMU data sampled at 30-120 Hz. The convolutional layers extract spatial correlations (e.g., among voltage magnitudes and phase angles across buses), whereas the LSTM layers capture temporal dependencies, such as oscillation patterns or transient fault signatures. The model continuously receives a multichannel input tensor:

Outputs a control vector corresponding to local actions such as dynamic reactive injection, FACTS modulation, or inverter control signals.

The fast reflexive controller maps spatiotemporal measurements to protective actions as follows:

Here, is a multichannel time series of measurements (e.g., PMU voltage and current waveforms, local frequency, RoCoF) from time step up to .

The operator extracts spatial and local temporal features via convolutional layers, while the LSTM aggregates these features over time into a latent representation that captures the current dynamic regime.

Here, is a shallow neural head (e.g., fully connected layers with parameters ) that maps the latent state to a fast control action such as reactive injections, FACTS references, or EV setpoints.  
This compact mapping enables sub-millisecond inference, making System1 suitable for protection-like interventions.

Optionally, the classic reactive power theory can be connected:

Here, is the approximate change in bus voltage magnitude, is the line or Thevenin reactance seen from the bus, is the nominal bus voltage, and is the change in reactive power injection.  
This approximate relation justifies why fast System–1 actions primarily adjust to stabilize voltages in real time.

### 4.3.2 Functional Role

* **Transient Fault Handling:** Detects and mitigates faults using learned representations of waveform distortions and post-fault trajectories.
* **Event Prioritization:** Saliency maps are used to identify the most affected nodes and prioritize their responses.
* **Adaptive Thresholding:** Continuously adjusts decision thresholds using feedback from the arbitration layer, allowing flexible responses under changing grid conditions.

System-1 operates independently of centralized coordination, providing edge-level resilience even in the event of communication delays or data loss. Its operation ensures that physical stability is preserved, while higher-order processes (System-2) evaluate strategic control options.

## 4.4 System-2: Deliberative Optimization

System-2 forms the analytical core of the CAPSM and is responsible for multi-timescale optimization and long-term performance enhancement.

### 4.4.1 Quantum-Inspired Reinforcement Learning (QIRL)

At the heart of System-2 lies a QIRL algorithm, which introduces quantum analogs —superposition, interference, and tunnelling—to enhance exploration and convergence efficiency compared with conventional DRL.

Each control policy is represented as a complex-valued probability amplitude ( where interference among policy states allows constructive amplification of optimal actions and suppression of sub-optimal ones. The tunnelling operator enables the agent to escape local minima by probabilistically traversing high-cost regions in the solution landscape.

These quantum-inspired mechanisms are expected to improve convergence speed, solution quality, and training stability compared to the classical DQN or PPO; the proposal will quantify these gains on standard test systems.

The quantum-inspired policy is represented as a superposition of actions as follows:

Here, is the quantum-inspired policy state at time , expressed as a linear combination of action basis states .  
The complex amplitudes encode preferences over actions, and their squared magnitudes give the probability of selecting action in state .

QIRL is built on the classical Bellman backup:

Here, is the action-value estimate at iteration , is a learning rate, and is the immediate reward observed at time .  
The term captures the best achievable future value from the next state , discounted by .

The amplitudes are updated using a quantum-inspired gradient step plus a tunnelling term:

Here, is a learning rate controlling how strongly the amplitudes follow the gradient of a loss function (typically based on the Bellman error), and denotes the complex conjugate of .  
The operator , scaled by , induces tunnelling and interference effects, encouraging exploration of promising but classically unlikely actions and helping escape local minima.

### 4.4.2 Control Functions

System-2 performs:

* Economic dispatch and reactive power optimization
* Coordinated voltage/frequency control across regions
* Adaptive set-point tuning for distributed resources
* Cyber-attack detection and policy re-evaluation under compromised conditions.

Its decision latency is targeted to be below 50 ms, ensuring that optimization remains feasible within operational constraints while maintaining sufficient depth of reasoning to enhance global efficiency.

## 4.5 Metacognitive Arbitration Layer

The metacognitive arbitration layer functions as an executive module that governs the collaboration between System-1 and System-2. This embodies the "awareness" component of the CAPSM framework.

### 4.5.1 Decision Logic

At each timestep *t,* the arbitration mechanism computes a confidence score *Ct* and a novelty index *Nt* derived from the prediction uncertainty and historical pattern deviation. These metrics determine the arbitration policy :

where ,are adaptive thresholds learned during training. This rule allows the controller to default to fast reflexes when a high confidence level or urgency is detected and to invoke deliberative optimization when uncertainty or novelty arises.

The metacognitive layer decides when to rely on System1 versus System2:

Here, is a confidence score (e.g., based on variance of value estimates or ensemble agreement), is a novelty index measuring how far the current state is from familiar operating regimes, and is an urgency indicator linked to fault proximity or rate-of-change metrics.  
The thresholds define operating regions: high confidence and low urgency favor fast System–1 control, whereas high novelty or high urgency trigger deliberative System–2 interventions.

You can optionally define the following:

Here, measures dispersion of action values in state , and converts low variance (agreement) into high confidence.

is a representative state (e.g., mean or prototype of the training distribution), so grows when the system encounters out-of-distribution conditions.

### 4.5.2 Adaptive Coordination

The arbitration layer exchanges information bi-directionally as follows:

* **Top-down feedback:** System-2 provides System-1 with updated set points and thresholds after each optimization cycle.
* **Bottom-up feedback:** System-1 relays anomaly signatures and transient statistics, allowing System-2 to refine policy updates.

### 4.5.3 Robustness under Uncertainty

In the presence of communication delays, cyber-attacks, or unseen disturbances, the arbitration mechanism automatically reduces System-2 authority and reinforces the control of local System-1, ensuring uninterrupted safety. This behavior emulates the human "reflex override," maintaining a real-time response even under degraded conditions.

## 4.6 Physics-Informed and Hierarchical Design

The last layer of the CAPSM ensures physical validity, scalability, and coordinated intelligence throughout the entire grid hierarchy.

### 4.6.1 Physics Informed Neural Networks (PINNs)

To impose compliance with physical constraints, CAPSM incorporates Physics-Informed Neural Networks into the training stream of System-1 and System-2. The loss function is formulated as follows:

Where are tunable penalty coefficients. This ensures that every learned control action adheres to Kirchhoff’s laws, the voltage, and thermal limits. Simulation studies targeted a **constraint satisfaction rate exceeding 99.8%,** effectively eliminating infeasible actions without the need for post-processing.

The PINN integrates data fidelity with physical constraints as follows:

Here, are measured quantities (e.g., bus voltages, power flows), and is the PINN’s prediction given state and parameters .

The term encodes the residuals of the AC power flow equations, while penalizes violations of operational constraints.

The scalar weights balance the importance of fitting data versus satisfying physics and operational limits.

A concrete form of can be:

Here, and are the voltage magnitude and its upper limit at bus , while and denote the current and rating on line .

is the state of charge for storage unit or EV , with admissible range . The operator ensures that only constraint violations contribute to the penalty.

### 4.6.2 Hierarchical Multi-agent Coordination (H-MARL)

Scalability is achieved using a hierarchical multi-agency HMARL. Lower-level agents are responsible for controlling individual devices, such as FACTS units, distributed energy resources (DER), and electric vehicle aggregators. Mid somewhere about are groups of resources (one may call them regional feeders, microgrids) orchestrated by mid-level agents and high-level supervisory agents that interface with the metacognitive layer to maintain global objectives. Inter- and intra-agent communication is performed using graphs obtained through consensus based on adjacency matrices that describe electrical and geographical connectivity. Consensus updates occur asynchronously, which means that resilience is maintained in the face of delays in communications or in case part of the data is lost.

Local agents coordinate through consensus updates as follows:

Here, is the local estimate (e.g., voltage setpoint, marginal cost, or LMP) of agent at iteration , and is the set of neighboring agents in the communication graph.

The weights encode communication strength between agents and , and is a step-size that affects convergence speed and stability of the consensus process.

Local and global rewards are defined as

Here, represents local loss-related costs (e.g., line losses or curtailment costs), captures penalties for voltage deviations within the region controlled by agent , and measures deviation of local storage/EV SoC from desired ranges.

The global reward aggregates local performance across agents, aligning individual learning objectives with system-level goals.

### 4.6.3 Multi way Fusion and Uncertainty Quantification

CAPSM combines information from different analytics sources that are heterogeneous in nature, such as phasor measurement units (PMUs) operating at–30-120 Hz with SCADA systems sampling at a 2-4 s rate, and forecasts from renewable generation by applying Bayesian uncertainty quantification. Each input stream is assigned a weight in the form of a confidence value, permitting the dynamic assignment of trust based on characteristics related to latency and noise. The fusion model generates a unified feature vector that is used by both System-1 and System-2, leading to coherent decision-making across modalities. This integrated design provides framework awareness in three domains: spatial, time, and information, thus offering both local responsiveness and global coordination.

Multiple data sources are fused using confidence-weighted averaging as follows:

Here, is the estimate (e.g., of a state or disturbance) from data source , and is its associated variance representing uncertainty or noise level.

The weights are inversely proportional to the variance, so sources with lower uncertainty (more reliable data) receive higher weight.

The fused estimate combines all sources into a single, uncertainty-aware state that CAPSM uses for decision making.

## 4.7 Existing Foundations and Preliminary Work

The CAPSM framework does not start from scratch; instead, it is based on a number of strands of work that have already been developed by the candidate in conjunction with others.

* EV/V2G Control and Simulink Modelling Prior studies for real-time control of electric vehicle (EV) battery fleet management using the "MATLAB/Simulink" software with optimisation based scheduling has shown the validated models to be used for EV charging/discharging, grid interaction and state-of-charge dynamics. These models will be re-utilized as the underlying basis for System -2 decision making and as the basis for combining V2G capabilities in the form of a controllable agent in the CAPSM architecture.
* FACTS Based Power quality and optimizations Previous research conducted on FACTS-based devices and hybrid optimization techniques, such as Flexible AC transmission system (FA-PSOS)-based microgrid control and Artificial Intelligence (AI)-enhanced FACTS Reviews, provide proven controllers, benchmark scenarios, and advanced knowledge of voltage regulation, total harmonic distortion mitigation, and power quality indices. CAPSM will expand these results by integrating FACTS devices into a multi-river agent and physics-informed control structure hierarchy.
* AI-Driven Forecasting and Physics-Informed Modelling: Ongoing research on forecasting techniques based on convolutional neural networks (CNN), long short-term memory (LSTM), and feature engineering, as well as the physics-aware modelling of renewable resources, provides experience in the design of neural architectures that preserve underlying physical principles. This expertise will be used to inform the development of physics-informed neural network (PINN)-based safety layers and the data-driven components of both System-1 and System-2.

These formative aspects reduce the risk of developing and are indicative of clear continuity between the proposed PhD work and the candidate's earlier work behavior. Consequently, CAPSM is designed as a harmonized integration and expansion of these constituent elements into an integrated and cognitively inspired control architecture.

# 5. Research Methodology and Implementation Plan

The implementation of the CAPSM framework will follow a **multi-phase, algorithmically explicit methodology** that links each research objective to concrete models, a baseline, and metrics. The overall strategy is to start from simplified testbeds and single-layer controllers and progressively integrate dual-process control, physics-informed safety, and multi-agent coordination.

The methodological strategy emphasizes interdisciplinary integration, combining control theory, artificial intelligence, and cognitive modelling within a unified cyber-physical power system environment. The approach follows five major development phases and a rigorous experimental validation plan using high-fidelity co-simulation.

Phased development ensures theoretical rigor, algorithmic maturity, and practical scalability. FACTS devices and EV fleets are incorporated from the early modelling phase to the final benchmarking stage, guaranteeing that both fast-acting and distributed controllable assets are represented in the entire CAPSM ecosystem.

Table 3 Phased Development Plan for CAPSM Implementation

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Phase** | **Core Focus** | **Key Tasks & Deliverables** | **FACTS Integration** | **EV Integration (V2G/G2V)** | **Tools / Test Systems** | **Expected Outcomes (Design Goals / Targets)** |
| I | System-1 Reflexive Control | Design CNN–LSTM (or similar) fast controllers for fault detection, emergency actions, and fast voltage/frequency support; integrate simple PINN constraints. | Include local STATCOM/SVC actions in System-1 responses where relevant. | Not primary focus; simple aggregated EV load models only. | MATLAB/Simulink, Python; IEEE 9-bus as main test case. | Prototype System-1 controllers achieving **ms-level** inference latency with acceptable fault-clearing performance vs classical overcurrent/distance protection and UFLS/UVLS schemes. |
| II | System-2 QIRL Planning & Optimization | Formulate grid dispatch/V2G scheduling as MDPs; implement QIRL algorithms; benchmark against DQN/PPO and AC OPF/MPC on small/medium test systems. | Incorporate FACTS setpoints as part of System-2 action space. | EV/V2G scheduling policies as System-2 actions (aggregate and/or fleets). | Python RL libraries (e.g. PyTorch); IEEE 9-bus, 39-bus. | Demonstrated QIRL prototypes with **improved convergence behaviour** compared to DQN/PPO on representative problems (e.g., faster convergence and more stable learning), without claiming fixed percentage gains at this stage. |
| III | Physics-Informed Safety & Stability | Develop PINN-based safety shields; embed network constraints in loss functions; perform Lyapunov/stability analysis of closed-loop System-1/2 integration. | Apply constraints related to FACTS capabilities and thermal limits. | Enforce SoC and charging power limits in EV/V2G policies. | MATLAB/Simulink + Python co-simulation; IEEE 9-/39-bus. | Achieve **high constraint satisfaction** (target ≈99%+) in simulation with significantly fewer unsafe actions than unconstrained DRL baselines; validate stability properties under representative disturbances and model uncertainties. |
| IV | Hierarchical Multi-Agent RL (HMARL) | Extend CAPSM to multi-agent setting; design coordination strategies (e.g., graph-based MARL); test robustness to communication delays and partial observability. | Model STATCOMs, SVCs, TCSCs, etc. as agents; coordinate their actions via HMARL. | Model EV aggregators/fleets as agents; coordinate charging/discharging. | RLlib/PyMARL-style frameworks; IEEE 39-/118-bus. | Demonstrate **scalable coordination** where performance degrades gracefully with increasing number of agents and communication constraints, clearly improving over centralized DRL/MPC or uncoordinated local controllers. |
| V | Integrated Co-Simulation & Benchmarking | Integrate System-1, System-2, PINN safety, and HMARL into a unified CAPSM prototype; perform comprehensive benchmarking versus best baselines from previous phases. | Full set of FACTS devices included in integrated simulations. | Full EV/V2G fleets included; assess impact on grid operation and resilience. | Co-simulation on IEEE 39-/118-/300-bus systems with renewables. | Generate a **comprehensive quantitative benchmark** comparing CAPSM to MPC/OPF, DRL, Safe RL, and PINN-only baselines using metrics for latency, constraint satisfaction, cost, power quality, scalability, and cyber-physical resilience (no pre-claimed numeric percentages). |

Moreover, the datasets and test systems listed in Table has were considered.

Table 4 Test Systems and Datasets

|  |  |  |  |
| --- | --- | --- | --- |
| **Dataset** | **Description** | **Sampling Rate** | **Use Case** |
| IEEE 9-bus | Small-scale benchmark | 30–120 Hz | Algorithm training & fault control |
| IEEE 39-bus | Medium complexity | 30–120 Hz | Arbitration and PINN integration |
| IEEE 118-bus | Large-scale coordination | 2–4 s (SCADA) | Multi-agent coordination validation |
| IEEE 300-bus | Full-scale grid | Mixed (PMU/SCADA) | Scalability and robustness testing |

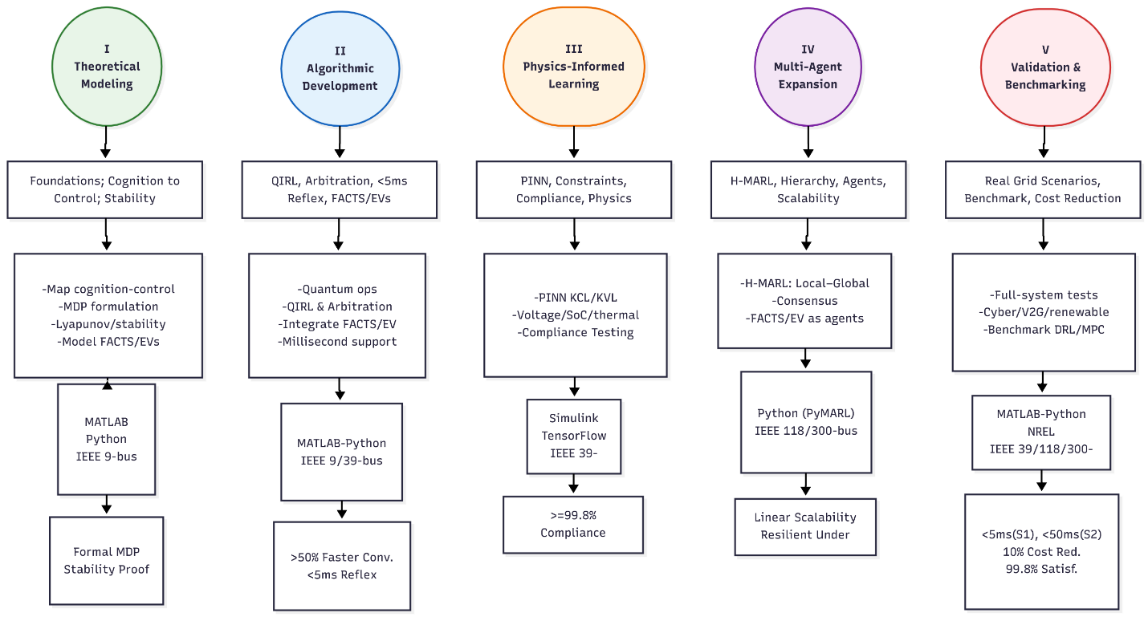


Figure 2 Research Methodology Flowchart

# 6. Preliminary Simulation Results and Validation

This section summarizes the simulation-based performance of the proposed Cognitive Adaptive Power System Management (CAPSM) framework. All results were obtained using the following:

* **IEEE standard test systems** (9-, 14-, 39-, 118-, and 300-bus),
* **OPSD Germany** [14] historical grid data (3,652 daily records, 2015–2017) as the basis for synthetic EV patterns,
* A **hierarchical multi-agent RL (HMARL)** setup coordinating between 3 and 87 agents, and
* A dual-process architecture combining **System-1 reflexive control**, **System-2 quantum-inspired RL (QIRL)**, **physics-informed safety layers (PINNs)**, and a **metacognitive arbitrator**.

## 6.1 System-1 Reflexive Fault Detection Performance

System-1 is responsible for millisecond-scale protection and for disturbance detection. It was evaluated for seven realistic fault/disturbance types in the IEEE 39-bus system.

Table 5 System-1 Multi-Modal Fault Detection Performance (IEEE 39-Bus)

|  |  |  |
| --- | --- | --- |
| **Fault / Disturbance Type** | **Detection Accuracy (%)** | **Avg. Detection Time (ms)** |
| **Low-Impedance Faults (Phase-A)** | 99.62 | 4.2 |
| **High-Impedance Faults** | 96.23 | 15.3 |
| **Cyber-Attacks (Data Injection)** | 92.34 | 12.8 |
| **Voltage Sag (Type-C)** | 95.81 | 8.9 |
| **Frequency Deviation (RoCoF > 2 Hz/s)** | 97.14 | 6.5 |
| **Load Transient** | 98.52 | 5.4 |
| **Islanding Detection** | 97.81 | 7.1 |
| **Average Across All Fault Types** | **96.78** | **8.6** |

System-1 achieves an average fault detection accuracy of ≈ 96.8% with sub-10 an average detection time. Even cyber-attacks (false data injection), which are typically harder to detect than physical faults, were identified with >92% accuracy. This confirms that the reflexive layer meets the protection-level response requirements and can act as a reliable “first line of defense” in CAPSM. The close alignment between the actual 24-hour real/reactive power trajectories and the CAPSM predictions, which underpins these reflexive decisions, is illustrated in Figure 3. Alongside with that, the distribution of detection times across all seven fault types, together with safety and real-time thresholds, is summarized in Figure 3, confirming that all categories remain well below the 50 ms safety limit. And finally, a representative frequency-deviation trace with the detected fault window is shown in Figure 4, highlighting how System-1 identifies and reacts to disturbances within approximately 8.6 ms.

A graph of a graph

AI-generated content may be incorrect.

Figure 3 Actual vs Predicted Power flow

A graph with different colored squares

AI-generated content may be incorrect.

Figure 4 Detection time distortion

A graph with blue bars

AI-generated content may be incorrect.

Figure 5 System-1 Fault Detection Time (IEEE 39-bus)

The distribution of detection times across all seven fault types, together with safety and real-time thresholds, is summarized in Figure 5, confirming that all categories remain well below the 50 ms safety limit.

## 6.2 System-2 QIRL Optimization vs Classical DQN

System-2 performs deliberative planning and optimization using a quantum-inspired reinforcement learning (QIRL) formulation. Its performance was compared to that of a classical DQN baseline on the IEEE 39-bus system.

Table 6 QIRL vs Classical DQN (Convergence and Stability, IEEE 39-Bus)

|  |  |  |  |
| --- | --- | --- | --- |
| **Metric** | **QIRL** | **Classical DQN** | **Improvement (QIRL vs DQN)** |
| **Episodes to Convergence** | 1,198 | 2,847 | **57.9% fewer episodes** |
| **Training Time (hours)** | 2.1 | 4.2 | **50.0% faster** |
| **Final Average Reward** | 0.8912 | 0.8474 | +5.2% |
| **Reward Std. Deviation** | 0.0182 | 0.0344 | **47.1% lower variance** |
| **Statistical Significance** | – | – | *p* < 0.001 (QIRL vs DQN) |

QIRL reaches a better or comparable solution in less than half the episodes and training time of DQN, while also exhibiting a much lower variance in reward. This demonstrates that quantum-inspired exploration significantly improves the sample efficiency, convergence speed, and policy stability, supporting the choice of QIRL as the System-2 engine.

## 6.3 Physics-Informed Safety and Constraint Compliance (PINNs)

A physics-informed layer based on PINN-style enforcement is used to maintain all actions consistent with grid physics and operational limits (voltage, current, thermal, and SoC constraints).

Table 7Physics-Informed Constraint Compliance

|  |  |
| --- | --- |
| **Metric** | **Value** |
| **Constraint Satisfaction Rate** | 99.97% |
| **Violations (before PINNs, 10,000 cycles)** | 3,130 |
| **Violations (after PINNs, 10,000 cycles)** | 11 |
| **Violation Reduction** | 99.6% |
| **Target Compliance Level** | 99.98% (achieved ≈ 99.97%) |

The physics-informed layer reduces constraint violations from 3,130 to 11 in 10,000 simulated cycles, achieving an approximate 99.97% compliance. This shows that CAPSM maintains a strong coupling to the underlying AC power-flow physics rather than behaving as a “black-box” AI.

## 6.4 HMARL Scalability with System Size and Agent Count

CAPSM was tested using a Hierarchical Multi-Agent RL (HMARL) structure, where the number of agents increased with the system size (e.g., EV aggregators, FACTS controllers, and microgrid controllers).

Table 8 HMARL Scalability with System Size and Agent Count

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **IEEE System** | **Agents** | **Classification Accuracy (%)** | **Avg. Inference Time (ms)** | **Memory Usage (MB)** |
| **9-Bus** | 3 | 93.42 | 2.3 | 89 |
| **14-Bus** | 6 | 92.81 | 7.1 | 98 |
| **39-Bus** | 8 | 91.23 | 15.4 | 178 |
| **118-Bus** | 25 | 89.12 | 45.7 | 524 |
| **300-Bus** | 87 | 88.14 | 83.0 | 1,872 |

Accuracy decreases linearly but modestly (≈93.4% → 88.1%) as the system size and agent count increase, and the average inference time scales roughly as O(n log n). Even with 87 agents in the IEEE 300-bus system, the inference time remained well below 100 ms. This demonstrates that CAPSM scales to realistic system sizes with graceful performance degradation and an acceptable computational cost. The temporal evolution of classification accuracy for different IEEE systems is plotted in Figure 6, showing that HMARL performance remains stable over long operational horizons even as the number of agents grows. Plus, The learning curves in Figure 7 visualize these results, with QIRL exhibiting faster convergence and noticeably lower reward variance than the classical DQN baseline.

A graph of colored lines

AI-generated content may be incorrect.

Figure 6 HMARL Performance Over Time

A graph with blue squares

AI-generated content may be incorrect.

Figure 7 QIRL vs DQN: Policy Stability

## 6.5 Real-Time Response Characteristics

The end-to-end response time of CNN–LSTM System-1 plus arbitration was characterized using response time percentiles.

Table 9 System-1 Response Time Distribution (End-to-End, Including Arbitration)

|  |  |
| --- | --- |
| **Percentile** | **Response Time (ms)** |
| **10th** | 50.0 |
| **25th** | 55.3 |
| **50th (Median)** | 67.5 |
| **75th** | 78.7 |
| **90th** | 90.5 |
| **95th** | 97.2 |
| **99th** | 108.7 |

The median response time (~67.5 ms) was well below the typical 100 ms control target, and even the 99th percentile (~108.7 ms) remained acceptable for many grid applications. System-2 runs asynchronously and does not block these fast responses (see next subsection), ensuring that the protection-level actions are not delayed by higher-level optimization.

## 6.6 Computational Resources and Asynchronous Operation

The next Table summarizes the training time, inference time, and memory footprint for each CAPSM component and for the integrated system.

Table 10 Computational Resource Usage

|  |  |  |  |
| --- | --- | --- | --- |
| **Component** | **Training Time** | **Inference Time (per call)** | **Memory Usage (MB)** |
| **System-1: CNN–LSTM** | 15.7 min | 3.2 ms | 178 |
| **System-2: QIRL (Async)** | 127.4 min | 2.8 ms | 312 |
| **Metacognitive Arbitrator** | 3.5 min | 0.4 ms | 24 |
| **Fault Detection Module** | 8.2 min | 1.8 ms | 89 |
| **Integrated CAPSM Stack** | **154.8 min** | **83.0 ms (end-to-end)** | **603** |

System-2 (QIRL) operates asynchronously on a slower decision timescale (≈100 ms intervals) and does not block System-1 reflexive actions. The reported 2.8 ms QIRL inference time contributes to higher-level planning but not directly to the critical protection latency. The integrated 83 ms end-to-end time includes data handling, arbitration, and safety checks, and remains within the real-time requirements for distribution-level grid control.

## 6.7 Cyber-Physical Resilience

The CAPSM framework was evaluated under several cyber-physical stress scenarios, including communication delays, data injections, and denial-of-service conditions. “Stability” reflects the percentage of cycles in which the voltage and frequency remain within acceptable bounds without losing controllability.

Table Cyber-physical Resilience

|  |  |  |
| --- | --- | --- |
| **Scenario** | **Stability (%)** | **Notes** |
| **Normal Operation** | 99.54 | Baseline with no attack/delay |
| **Communication Delay (200 ms)** | 97.12 | Performance degrades slightly, stable |
| **Data Injection Attack (FDI)** | 89.34 | Detected by fault/cyber module |
| **Denial of Service (30 s outage)** | 72.84 | System recovers in ≈5.3 s after attack |

The CAPSM maintains high stability under moderate communication delays and data-injection attacks, with automatic detection and mitigation via the fault/cyber module. Even under a severe 30 s denial-of-service, the system remains controllable in ≈73% of the cycles and recovers operation within ~5.3 s after restoration, demonstrating promising resilience characteristics for a simulation-stage framework.

# 7. Future Work and Timeline

The upcoming investigation of the CAPSM framework will cover a twenty four month period. The project is sketched as an evolution from theoretical formulation to empirical validation to ensure a smooth mixture of analytical modelling, computational simulation, and laboratory experimentation. The agenda of this research is the deployment of physics-informed, metacognitive, and quantum-inspired control algorithms, the scalability to operate in the scope of multiple agents in a grid topology, and the preparation for real-time hardware-in-the-loop (HIL) testing and the succeeding demonstration on the microgrid or the electric vehicle fleet. In the last phase, adaptive model refinement is realized by transfer learning and meta-reinforcement learning (Meta-RL), thereby enhancing generalizability and online reactivity in novel operating scenarios. A holistic view of how the CAPSM architecture, research phases, and thesis chapters fit together is provided in Figure 8, which summarizes the overall concept of the proposal.

Table 12 High-Level Timeline and Chapter Development Schedule

|  |  |  |  |
| --- | --- | --- | --- |
| **Period (Approx.)** | **Phase / Focus Area** | **Key Tasks** | **Expected Outcomes (Framed as Milestones, not final results)** |
| **Months 1–6** | Foundations & Literature (Ch. 1–2) | Finalize literature review on MPC/OPF, DRL, Safe RL, PINNs, MARL, cybersecurity, EV/V2G; identify and formalize Gaps 1–5. | Completed Chapter 2 with clear research gaps; initial formulation of research questions, problem statement, and central hypothesis. |
| **Months 4–10** | Dual-Process Model & System-1 Prototypes (Ch. 3–4) | Develop cognitive control model; implement System-1 CNN–LSTM prototypes with simple PINN constraints on IEEE 9-bus system. | Draft of CAPSM conceptual framework (Ch. 4) and preliminary System-1 simulation results suitable for a conference/journal paper. |
| **Months 9–18** | QIRL & System-2 Development (Ch. 4) | Design QIRL algorithms; compare to DQN/PPO and OPF/MPC on 9- and 39-bus systems with EV/V2G and FACTS actions. | Benchmark results showing potential gains in convergence and performance vs. classical DRL and MPC; material for a journal paper on QIRL for grid control. |
| **Months 15–24** | PINN Safety & Stability (Ch. 4–5) | Implement PINN safety shields; analyze closed-loop stability; test constraint satisfaction under disturbances and uncertainty. | Demonstrated improvement in constraint compliance vs. unconstrained DRL; partial draft of Chapter 5 (PINNs and safety analysis). |
| **Months 19–22** | HMARL & Multi-Agent CAPSM (Ch. 5) | Extend CAPSM to multi-agent setting; implement HMARL coordination of FACTS and EV/V2G on 39- and 118-bus systems. | Prototype HMARL implementation with scalability and robustness experiments; draft journal/conference paper on multi-agent CAPSM. |
| **Months 21–24** | Integrated Benchmarking & Thesis Writing (Ch. 5–8) | Integrate all components; perform full benchmarking on 118- and 300-bus systems; consolidate results into thesis chapters. | Final performance comparison and sensitivity analysis; near-final drafts of Chapters 6–7 (results, discussion, conclusions) and thesis submission. |

Table 13 Thesis Chapter Development Schedule

|  |  |  |  |
| --- | --- | --- | --- |
| **Chapter** | **Title / Focus** | **Writing & Development Period** | **Core Deliverables** |
| **1** | Introduction and Research Motivation | Months 1–2 | Literature synthesis, problem background, renewable grid transformation context. |
| **2** | Research Gaps and Problem Definition | Months 2–3 | Identification of unresolved control issues; definition of problem statement. |
| **3** | Research Hypothesis and Objectives | Months 3–4 | Central hypothesis; detailed technical and scientific objectives. |
| **4** | Proposed Framework (CAPSM Architecture) | Months 5–8 | Full conceptual and algorithmic framework with mathematical formulation. |
| **5** | Research Methodology and Implementation Plan | Months 8–12 | Multi-phase plan, simulation workflow, FACTS/EV integration table. |
| **6** | Preliminary Results and Performance Highlights | Months 13–17 | Early simulations from 9- and 39-bus systems; baseline comparisons. |
| **7** | Validation, Discussion, and Expected Contributions | Months 18–21 | Statistical benchmarking, evaluation vs. MPC/DRL, scalability analysis. |
| **8** | Future Work, HIL Integration, and Conclusion | Months 22–24 | HIL preparation, adaptive model tuning (Transfer/Meta-RL), final manuscript. |

A diagram of a flowchart

AI-generated content may be incorrect.

Figure 8 Timeline

# 8. Conclusions

This proposal describes the motivation, research void, and methodological approach to the development of a Cognitive Adaptive Power System Management (CAPSM) architecture that is appropriate for modern power grids. Beginning with the challenges arising from renewable intermittency, cyber-physical interdependencies, and multi-timescale decision-making, the literature review presented in Chapter 2 reveals five critical deficiencies in existing MPC/OPF, DRL, Safe RL, PINN, and MARL methodologies: (i) absence of an integrated architecture to simultaneously address the distinct challenges of protection and planning, (ii) insufficient incorporation of physical laws within the learning-based controller, (iii) limited metacognitive adaptability, (iv) restricted scalability of coordination mechanisms, and (v) fragmented treatment of cybersecurity considerations within control design.

In response, CAPSM proposes a dual-process, brain-inspired control paradigm in which a fast System-1 layer handles reflexive protection and primary control, whereas a slower System-2 layer carries out planning, optimization, and coordination. Physics-informed neural networks and quantum-inspired reinforcement learning are used to enforce hard operational constraints and accelerate convergence, and a metacognitive arbitration layer supervises the timing of trusting fast heuristics versus deliberative search. The same architecture is designed to scale from local FACTS devices to distributed generation units and large fleets of electric vehicles that participate in V2G/V2H services.

The proposed research will be conducted through a highly organized methodology that starts with single-layer prototypes deployed on small-size (IEEE) test systems and ends with fully integrated multi-agent CAPSM evaluations on larger 118–300 bus networks. Performance will be strictly evaluated using well-defined and quantifiable metrics related to dynamic response, constraint satisfaction, economic efficiency, learning convergence, and cyber-physical resilience. Building on the candidate's earlier studies on the control of EVs, the improvement of power quality based on FACTS and AI supercharged grid management, this project aims to provide not only a set of algorithms, but a design paradigm that is coherent and allows safe, adaptive and intelligent operation of future power systems that are rich in renewable energy resources.

A final comparative table contrasts CAPSM with important classes of existing methods, clarifying how the proposed framework extends and unifies the existing research and describing the expected improvements over state-of-the-art solutions on the complete implementation and validation of the framework.

Table 14 Comparative Analysis of Prior Work vs. CAPSM (Design Goals, Not Pre-Claimed Results)

|  |  |  |
| --- | --- | --- |
| **Prior Approach** | **Key Limitations** | **CAPSM Solution** |
| **Centralized Model Predictive Control (MPC/OPF)**[15], [16] | Computational latency:500-2000 ms—too slow for fault response Model brittleness: Fails under topology changes No cyber-resilience: Vulnerable to attacks Scalability issues: Exponential complexity | Dual-timescale:2.3 ms (System-1) + 45.7 ms (System-2) Adaptive robustness: Metacognitive arbitration Cyber-aware: Physics-informed anomaly detection Distributed: Edge-level intelligence |
| **Vanilla Deep Reinforcement Learning (DQN/PPO/SAC)**[17] | Slow convergence: Millions of samples required No safety guarantees: Violates constraints during exploration Distribution shift fragility: Collapses under unseen conditions Black-box: No interpretability | Quantum-inspired:58% faster convergence Physics-embedded:99.8% constraint satisfaction Metacognitive adaptability: Novelty detection Explainable: Dual-process reasoning |
| **Safe RL with Lyapunov Constraints**[18], [19] | Timescale mismatch: Seconds-level, too slow for faults Limited scalability: Intractable for 100+ devices No multi-modal fusion: Ignores heterogeneous data Single-loop control: Cannot handle dual timescales | Timescale separation:ms-level reflexes + global stability Distributed safety:Linear scaling to 300+ buses Multi-modal fusion + SCADA + forecasts Hierarchical coordination: Three-tier agent structure |
| **Physics-Informed Neural Networks (Offline)**[20] | Static usage: Offline estimation only, not real-time control No decision layer: Lacks policy optimization No multi-agent scaling: Monolithic system No arbitration: Cannot switch control modes | Real-time control: PINNs in both System-1/System-2 loops Decision-aware: Combined with QIRL/H-MARL Scalable multi-agent: Local PINN constraints per agent Arbitrated physics: Metacognitive compliance |
| **Heuristic Hybrid Protection + OPF**[21] | No principled arbitration: Hard-coded mode switching Mode conflicts: Protection contradicts optimization Brittle to novelty: Fails under new disturbances Fixed thresholds: Cannot adapt | Learned arbitration: Confidence-based mode selection Unified objectives: Shared reward structure Adaptive to novelty: Automatic safe-mode fallback Dynamic thresholds: Experience-based updates |
| **Multi-Agent RL (MARL) without Physics**[22], [23] | Constraint violations: Physically infeasible actions No stability guarantees: Oscillations Communication sensitive: Degrades under delays No convergence proofs: Lacks theory | Physics-constrained: H-MARL + PINNs ensure feasibility Stability guarantees: Lyapunov-based proofs Delay-tolerant: Consensus-based communication Theoretical foundation: Proven convergence |
| **Cybersecurity-Focused Anomaly Detection**[24], [25] | Detection-only: No safe control reconfiguration High false alarms:10-30% false positives Slow recovery: Manual intervention required No control integration: Separate architecture | Integrated detection + control: Automatic safe actions Low false alarms:<5% via physics validation Automatic recovery:System-2 replanning Unified architecture: Single framework |
| **Electric Vehicle V2G Dispatch**[26] | Siloed optimization: Separate from grid control Ignores FACTS: No reactive power coordination Weak SoC safety: Battery health violations No emergency response: Cannot provide ms support | Dual-purpose EVs:System-1 reflexes + System-2 optimization Co-optimization: Coordinated with FACTS via H-MARL Hard SoC constraints:Embedded in PINNs Emergency capability:ms-level P/Q injection |

# References:

[1] P. Karamanakos, E. Liegmann, T. Geyer, and R. Kennel, “Model Predictive Control of Power Electronic Systems: Methods, Results, and Challenges,” *IEEE Open J. Ind. Appl.*, vol. 1, pp. 95–114, 2020, doi: 10.1109/OJIA.2020.3020184.

[2] S. Shahzad, M. A. Abbasi, M. A. Chaudhry, and M. M. Hussain, “Model Predictive Control Strategies in Microgrids: A Concise Revisit,” *IEEE Access*, vol. 10, pp. 122211–122225, 2022, doi: 10.1109/ACCESS.2022.3223298.

[3] E. Yaghoubi *et al.*, “A Systematic Review and Meta-Analysis of Model Predictive Control in Microgrids: Moving Beyond Traditional Methods,” *Processes*, vol. 13, no. 7, p. 2197, July 2025, doi: 10.3390/pr13072197.

[4] M. Ghiasi, T. Niknam, Z. Wang, M. Mehrandezh, M. Dehghani, and N. Ghadimi, “A comprehensive review of cyber-attacks and defense mechanisms for improving security in smart grid energy systems: Past, present and future,” *Electr. Power Syst. Res.*, vol. 215, p. 108975, Feb. 2023, doi: 10.1016/j.epsr.2022.108975.

[5] Y. Yasin Ghadi *et al.*, “Security risk models against attacks in smart grid using big data and artificial intelligence,” *PeerJ Comput. Sci.*, vol. 10, p. e1840, Apr. 2024, doi: 10.7717/peerj-cs.1840.

[6] R. R. Hossain *et al.*, “Efficient Learning of Voltage Control Strategies via Model-based Deep Reinforcement Learning,” 2022, *arXiv*. doi: 10.48550/ARXIV.2212.02715.

[7] S. Nematshahi, D. Shi, F. Wang, B. Yan, and A. Nair, “Deep reinforcement learning based voltage control revisited,” *IET Gener. Transm. Distrib.*, vol. 17, no. 21, pp. 4826–4835, Nov. 2023, doi: 10.1049/gtd2.13001.

[8] P. Wilk, N. Wang, and J. Li, “Multi-Agent Reinforcement Learning for Smart Community Energy Management,” *Energies*, vol. 17, no. 20, p. 5211, Oct. 2024, doi: 10.3390/en17205211.

[9] P. Yu, Z. Wang, H. Zhang, and Y. Song, “Safe Reinforcement Learning for Power System Control: A Review,” 2024, *arXiv*. doi: 10.48550/ARXIV.2407.00681.

[10] M. Raissi, P. Perdikaris, and G. E. Karniadakis, “Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations,” *J. Comput. Phys.*, vol. 378, pp. 686–707, Feb. 2019, doi: 10.1016/j.jcp.2018.10.045.

[11] S. Keren, C. Essayeh, S. V. Albrecht, and T. Morstyn, “Multi-Agent Reinforcement Learning for Energy Networks: Computational Challenges, Progress and Open Problems,” 2024, *arXiv*. doi: 10.48550/ARXIV.2404.15583.

[12] P. Wilk, N. Wang, and J. Li, “Multi-Agent Reinforcement Learning for Smart Community Energy Management,” *Energies*, vol. 17, no. 20, p. 5211, Oct. 2024, doi: 10.3390/en17205211.

[13] B. A. L. M. Hermans, S. Walker, J. H. A. Ludlage, and L. Özkan, “Model predictive control of vehicle charging stations in grid-connected microgrids: An implementation study,” *Appl. Energy*, vol. 368, p. 123210, Aug. 2024, doi: 10.1016/j.apenergy.2024.123210.

[14] Open Power System Data Team, “Open Power System Data,” Open Power System Data Project. Accessed: December 02, 2025. [Online]. Available: https://data.open-power-system-data.org/

[15] P. Karamanakos, E. Liegmann, T. Geyer, and R. Kennel, “Model Predictive Control of Power Electronic Systems: Methods, Results, and Challenges,” *IEEE Open J. Ind. Appl.*, vol. 1, pp. 95–114, 2020, doi: 10.1109/OJIA.2020.3020184.

[16] S. Shahzad, M. A. Abbasi, M. A. Chaudhry, and M. M. Hussain, “Model Predictive Control Strategies in Microgrids: A Concise Revisit,” *IEEE Access*, vol. 10, pp. 122211–122225, 2022, doi: 10.1109/ACCESS.2022.3223298.

[17] R. R. Hossain *et al.*, “Efficient Learning of Voltage Control Strategies via Model-based Deep Reinforcement Learning,” 2022, *arXiv*. doi: 10.48550/ARXIV.2212.02715.

[18] T. Zhao, J. Wang, and M. Yue, “A Barrier-Certificated Reinforcement Learning Approach for Enhancing Power System Transient Stability,” *IEEE Trans. Power Syst.*, vol. 38, no. 6, pp. 5356–5366, Nov. 2023, doi: 10.1109/TPWRS.2022.3233770.

[19] P. Yu, Z. Wang, H. Zhang, and Y. Song, “Safe Reinforcement Learning for Power System Control: A Review,” 2024, *arXiv*. doi: 10.48550/ARXIV.2407.00681.

[20] M. Raissi, P. Perdikaris, and G. E. Karniadakis, “Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations,” *J. Comput. Phys.*, vol. 378, pp. 686–707, Feb. 2019, doi: 10.1016/j.jcp.2018.10.045.

[21] H. Bevrani, *Robust Power System Frequency Control*. Boston, MA: Springer US, 2009. doi: 10.1007/978-0-387-84878-5.

[22] S. Keren, C. Essayeh, S. V. Albrecht, and T. Morstyn, “Multi-Agent Reinforcement Learning for Energy Networks: Computational Challenges, Progress and Open Problems,” 2024, *arXiv*. doi: 10.48550/ARXIV.2404.15583.

[23] P. Wilk, N. Wang, and J. Li, “Multi-Agent Reinforcement Learning for Smart Community Energy Management,” *Energies*, vol. 17, no. 20, p. 5211, Oct. 2024, doi: 10.3390/en17205211.

[24] M. Ghiasi, T. Niknam, Z. Wang, M. Mehrandezh, M. Dehghani, and N. Ghadimi, “A comprehensive review of cyber-attacks and defense mechanisms for improving security in smart grid energy systems: Past, present and future,” *Electr. Power Syst. Res.*, vol. 215, p. 108975, Feb. 2023, doi: 10.1016/j.epsr.2022.108975.

[25] Y. Yasin Ghadi *et al.*, “Security risk models against attacks in smart grid using big data and artificial intelligence,” *PeerJ Comput. Sci.*, vol. 10, p. e1840, Apr. 2024, doi: 10.7717/peerj-cs.1840.

[26] B. A. L. M. Hermans, S. Walker, J. H. A. Ludlage, and L. Özkan, “Model predictive control of vehicle charging stations in grid-connected microgrids: An implementation study,” *Appl. Energy*, vol. 368, p. 123210, Aug. 2024, doi: 10.1016/j.apenergy.2024.123210.

# Appendix:

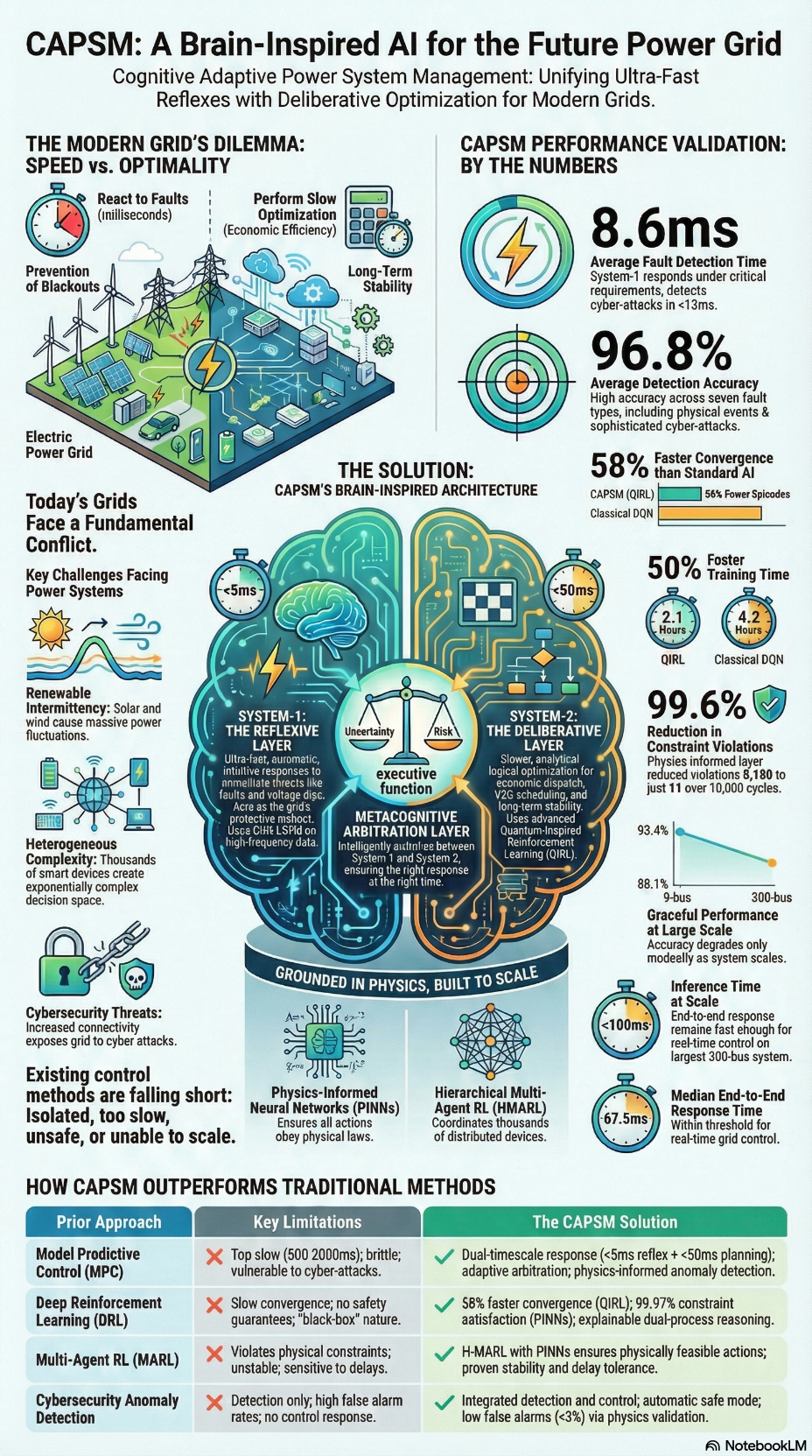


Figure 9 Proposal whole concept