Reinforcement Learning Meets the Power Grid: A Contemporary Survey with Emphasis on Safety and Multi-agent Challenges

Other titles in Foundations and Trends® in Electric Energy Systems

Voltage Stability of Power Systems
Saikat Chakrabarti
ISBN: 978-1-63828-496-3

Economics of Grid-Supported Electric Power Markets: A Fundamental Reconsideration
Leigh Tesfatsion
ISBN: 978-1-63828-428-4

Toward DyMonDS Framework for Resilient and Clean Electricity Services: The Puerto Rico Study
Marija D. Ilić, Rupamathi Jaddivada and Laurențiu Lucian Anton
ISBN: 978-1-63828-414-7

Operational Planning for Emerging Distribution Systems: A Unique Perspective on Grid Expansion
Anna Stuhlmacher, Chee-Wooi Ten, Lawrence Dilworth and Yachen Tang
ISBN: 978-1-63828-300-3

Distributed Optimization for the DER-Rich Electric Power Grid
Jannatul Adan and Anurag K. Srivastava
ISBN: 978-1-63828-292-1

Distribution System Optimization to Manage Distributed Energy Resources (DERs) for Grid Services
Anamika Dubey and Sumit Paudyal
ISBN: 978-1-63828-188-7

Reinforcement Learning Meets the Power Grid: A Contemporary Survey with Emphasis on Safety and Multi-agent Challenges

Ming Jin
Virginia Tech
jinming@vt.edu

the essence of knowledge
Boston — Delft

Foundations and Trends® in Electric Energy Systems

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
United States
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is

M. Jin. *Reinforcement Learning Meets the Power Grid: A Contemporary Survey with Emphasis on Safety and Multi-agent Challenges*. Foundations and Trends® in Electric Energy Systems, vol. 8, no. 3-4, pp. 169–316, 2025.

ISBN: 978-1-63828-577-9

Foundations and Trends® in Electric Energy Systems
Volume 8, Issue 3-4, 2025

Editorial Board

Editorial Scope

Foundations and Trends® in Electric Energy Systems publishes survey and tutorial articles in the following topics:

- Advances in power dispatch
- Demand-side and grid scale data analytics
- Design and optimization of electric services
- Distributed control and optimization of distribution networks
- Distributed sensing for the grid
- Distribution systems
- Fault location and service restoration
- Integration of physics-based and data-driven modeling of future electric energy systems
- Integration of Power electronics, Networked FACTS
- Integration of renewable energy sources
- Interdependence of power system operations and planning and the electricity markets
- Microgrids
- Modern grid architecture
- Power system analysis and computing
- Power system dynamics
- Power system operation
- Power system planning
- Power system reliability
- Power system transients
- Security and privacy
- Stability and control for the whole multi-layer (granulated) network with new load models (to include storage, DR, EVs) and new generation
- System protection and control
- The new stability guidelines and control structures for supporting high penetration of renewables (>50%)
- Uncertainty quantification for the grid
- System impacts of HVDC

Information for Librarians

Foundations and Trends® in Electric Energy Systems, 2025, Volume 8, 4 issues. ISSN paper version 2332-6557. ISSN online version 2332-6565. Also available as a combined paper and online subscription.

Contents

Reinforcement Learning Meets the Power Grid: A Contemporary Survey with Emphasis on Safety and Multi-agent Challenges

Ming Jin

Department of Electrical and Computer Engineering, Virginia Tech, USA; jinming@vt.edu

ABSTRACT

Modern power systems face increasing challenges from renewable energy integration, distributed resources, and complex operational requirements. This survey examines Safe Reinforcement Learning (Safe RL) as a framework for maintaining reliable power system operation while optimizing performance. We review both model-free and model-based approaches, analyzing how different safety constraints and architectures can be implemented in practice. The survey explores multi-agent frameworks for coordinated control in distributed settings and examines runtime assurance methods that provide formal safety guarantees. Applications span various timescales, from frequency regulation to demand management, with different safety requirements and operational contexts. Through analysis of current simulation environments and practical implementations, we identify remaining challenges in scaling safe RL to large power systems, handling uncertainty, and integration with existing infrastructure.

Ming Jin (2025), "Reinforcement Learning Meets the Power Grid: A Contemporary Survey with Emphasis on Safety and Multi-agent Challenges", Foundations and Trends® in Electric Energy Systems: Vol. 8, No. 3-4, pp 169–316. DOI: 10.1561/3100000043.

1

Introduction

1.1 Modern Power System Challenges

The ongoing evolution of power systems presents a multifaceted challenge: *ensuring safe and reliable operation amidst a dynamic and uncertain environment*. This necessitates not only achieving performance objectives but also adhering to diverse constraints encompassing operational limits, regulatory compliance, and environmental goals.

Key challenges in modern power systems include:

- *Uncertainty and Variability Challenges*: The integration of intermittent renewables, volatile demand, climate change impacts, and market price fluctuations introduce significant uncertainty, making it challenging to predict and manage power supply and demand.
- *Complexity and Scale Challenges*: Decentralization, diverse technologies (e.g., electric vehicles), interconnected grids, and increased digital reliance create a complex and multifaceted power system requiring sophisticated coordination and holistic management.
- *Reliability and Resilience Challenges*: Reduced system inertia and the increasing frequency of natural disasters necessitate rapid

response capabilities and robust recovery strategies to ensure grid stability and continuity of service.

- *Environmental and Regulatory Challenges*: Balancing the stringent environmental goals (e.g., reduce carbon emissions) with system stability and navigating complex regulations is crucial for ensuring a sustainable and resilient energy future.

1.2 Overview of Safe RL Applications in Power Systems

RL, with its adaptive learning capabilities, ability to handle high-dimensional spaces, and sequential decision-making framework, aligns well with the dynamic and complex nature of modern power grids. Furthermore, RL in its multi-agent form is essential for addressing the increasing complexity and scale of power systems, allowing for effective coordination of distributed energy resources, including electric vehicles, and management of intricate grid topologies. By learning from real-time interactions with the environment and optimizing for long-term rewards, RL has the potential to develop sophisticated control policies that outperform traditional rule-based systems. This could lead to more autonomous, efficient, and resilient power system operations (Figure 1.1).

Table 1.1 outlines seven critical power system applications where safe RL shows promise. These applications span economic optimization (Optimal Power Flow (OPF), energy management), system stability (frequency control, Volt-Var Control (VVC)), and reliability (Critical Load Restoration (CLR), Distribution Network Reconfiguration (DNR)). The need to handle uncertainties from renewable energy sources (RES) and variable demands, alongside the inherent complexity of power systems, makes these applications well-suited for RL approaches.

Safety constraints are application-specific, reflecting diverse objectives and operational contexts. For instance, OPF prioritizes voltage and line flow limits, while frequency control focuses on frequency stability and Rate of Change of Frequency (RoCoF). These constraints define the boundaries for RL agent operation. Across all applications, violations of safety constraints could lead to equipment damage, system instability, regulatory non-compliance, or service disruptions.

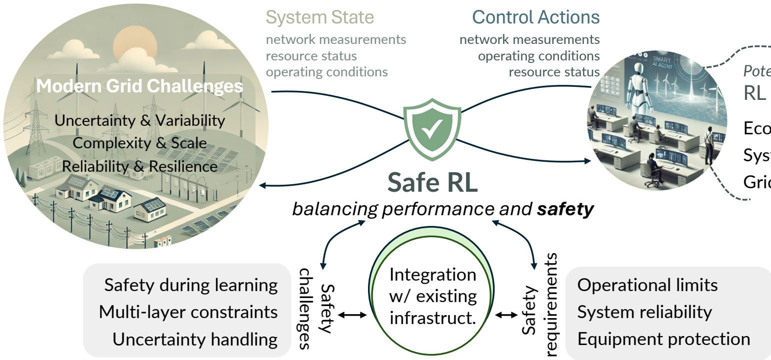

Figure 1.1: Safe RL for modern power systems. The framework processes system states, which may comprise of network measurements (voltage magnitude/angles at buses, active/reactive power flows, system frequency), resource status (generation outputs, storage SOC, RES availability), and operating conditions (load patterns, network topology, equipment status). The Safe RL module needs to address key challenges including safety during learning, multi-layered constraints, and uncertainty handling. It determines control actions, e.g., economic operations (generator setpoints, storage schedules), grid stability (AGC signals, reactive power control, tap changes), and emergency control (load restoration, network reconfiguration), while adhering to safety constraints (e.g., voltage bounds 0.95-1.05 pu, frequency ranges 59.8-60.2 Hz, thermal limits, stability margins, N-1 security). This can be typically implemented as either a safety layer on top of RL or as a simplex architecture (see Section 5). This enables various RL applications spanning economic optimization (OPF, energy management), system stability (frequency control, VVC), and grid reliability (load restoration, network reconfiguration).

The diversity of decision variables (continuous, discrete, mixed) across applications influences RL algorithm selection. Additionally, applications span transmission and distribution levels, each with unique challenges: transmission-level applications (e.g., OPF) often involve larger-scale considerations, while distribution-level applications (e.g., VVC) face higher uncertainty due to limited information.

The diverse requirements across power system applications create a complex landscape for safe RL. Real-time transmission-level applications (e.g., frequency control) necessitate rapid decision-making with continuous variables under strict safety constraints, whereas distribution-level applications (e.g., DNR) allow for more computational time but involve discrete decisions and complex network topology constraints.

Table 1.1: Overview of power system applications for safe RL. V: Voltage, Q: Reactive Power. Time scales – RT: Real-time, S: Short-term (minutes to hours), M: Medium-term (hours to days). System levels – D: Distribution, T: Transmission. Action types - C: Continuous, D: Discrete, M-D/C: Mixed Discrete-Continuous.

	Objective	**Challenges**	**Why RL?**	**Safety**	**Features**
OPF	min. costs w/ constraints	RES uncertainty, fast computation	fast decisions, adaptability	V limits, line flows, oper. limits	S; M-D/C; T
Energy Mgmt.	balance supply/demand, min. costs	gen./ demand uncertainty, prices	adapt, learn strategies	grid stability, V levels	S-M; M-D/C; D/T
Freq. control	maintain freq. in range	uncertainties, RES dynamics	adapt to rapid changes	freq. stability, RoCoF	RT; C; T/D
VVC	manage V profiles, Q flow	V fluctuations, rev. power flow	coord. control w/o full info	V range, device limits	S; M-D/C; D
CLR	restore critical loads	multi-step decisions, DER uncertainty	handle complexity, uncertainty	power flow constraints, stability	S; M-D/C; D
DNR	optimize feeder topology	incomplete info, computation	RT application, handle uncertainty	radial, V/f stability	M; D; D
EV charging	optimize charging schedules	variable demand, RES integration	adapt to changing conditions	grid stability, V levels	S-M; C; D

1.3 Safe RL: Bridging the Gap to Power System Applications

The primary challenge of applying standard RL to power systems is ensuring safety, given the potential for catastrophic consequences in this critical infrastructure, ranging from equipment damage and financial losses to life-threatening blackouts. Standard RL faces limitations in addressing power system safety due to:

- *Safety During Decision-Making*: One of the foremost challenges identified is ensuring safety during the learning and decision-making processes. As power systems operate under dynamic conditions, Safe RL (SRL) algorithms must guarantee safe performance while adapting to real-time changes in the environment. Failure to maintain safety can lead to critical system failures, emphasizing the need for robust safety mechanisms in RL applications

- *Multi-layered & Dynamic Constraints*: Power system constraints span various levels (e.g., physical equipment limitations, system-level stability requirements, regulatory rules) and can change over time, making comprehensive handling difficult for standard RL.
- *Handling Uncertainties*: Another significant challenge is managing the uncertainties that are prevalent in power system operations, such as fluctuations in demand and variability in renewable energy sources. SRL techniques must be capable of effectively coping with these uncertainties to make reliable predictions and decisions. Studies have indicated that existing algorithms often struggle with robustness in uncertain environments, impacting their practical applicability.
- *Complex Safety-Performance Trade-offs*: Finding the right balance between safety and optimal performance poses an ongoing challenge. Overly conservative safety constraints can hinder the efficiency of power systems, while inadequate focus on safety may lead to operational risks. Balancing these competing priorities is essential for the successful application of SRL.
- *Scalability & Uncertainty*: Ensuring system-wide safety while coordinating numerous distributed resources and handling rare events under uncertainty poses a significant challenge.
- *Integration with Existing Infrastructure*: Integrating SRL approaches with existing power system infrastructure also presents challenges. Many current systems were not designed with advanced machine learning strategies in mind, making it difficult to implement SRL solutions directly. The need for seamless integration is critical to harness the benefits of RL without disrupting existing operations.

These limitations reveal persistent challenges that must be addressed by SRL to ensure safe, efficient, and reliable operations in evolving energy landscapes. It's not merely an incremental improvement but a crucial adaptation designed to address the unique safety challenges of power systems. By prioritizing safety from the outset, SRL ensures operational safety, regulatory compliance, and risk mitigation, thus helping pave the way for wider adoption of RL in this critical domain.

1.4 Safe RL Formulations for Power Systems

RL is formulated as Markov Decision Processes (MDPs), defined by the tuple $\mathcal{M} \coloneqq \langle \mathcal{S}, \mathcal{A}, \mathbb{P}, r, \gamma, \rho \rangle$, where $\mathcal{S}$ represents the state space, $\mathcal{A}$ the action space, $\mathbb{P} : \mathcal{S} \times \mathcal{A} \to \Delta(\mathcal{S})$ the transition function governing state transitions based on actions, $r : \mathcal{S} \times \mathcal{A} \to \mathbb{R}$ the reward function quantifying the desirability of state-action pairs, $\gamma \in [0, 1)$ the discount factor weighing future rewards, and $\rho \in \Delta(\mathcal{S})$ the initial state distribution. In power systems, as illustrated in Figure 1.1, $s_t \in \mathcal{S}$ includes system states such as network measurements, resources status and operating conditions, while $a_t \in \mathcal{A}$ could represent control actions such as economic operations, grid stability and emergency control.

A common performance measure is the expected cumulative reward discounted over the infinite horizon:

$$J_r(\pi) \coloneqq \mathbb{E}_\pi[\sum_{t=0}^{\infty} \gamma^t r(s_t, a_t)] \tag{1.1}$$

Here, $\mathbb{E}_\pi[\cdot]$ denotes expectation over trajectory $\tau = (s_0, a_0, s_1, ...)$ under policy π and stochastic transition dynamics $\mathbb{P}$: $s_0 \sim \rho$, $a_t \sim \pi(\cdot|s_t), s_{t+1} \sim \mathbb{P}(\cdot|s_t, a_t)$. To make the dependence on state and action explicit, we express the on-policy value function as $V_r^\pi(s) \coloneqq \mathbb{E}_\pi[\sum_{t=0}^{\infty} \gamma^t r(s_t, a_t)|s_0 = s]$, the on-policy action-value function (or Q function) as $Q_r^\pi(s, a) \coloneqq \mathbb{E}_\pi[\sum_{t=0}^{\infty} \gamma^t r(s_t, a_t)|s_0 = s, a_0 = a]$, and the advantage function as $A_r^\pi(s, a) \coloneqq Q_r^\pi(s, a) - V_r^\pi(s)$. Another often used quantity is the discounted future state distribution (or occupancy measure), $d^\pi(s) \coloneqq (1 - \gamma) \sum_{t=0}^{\infty} \gamma^t \mathbb{P}(s_t = s|\pi)$, which allows us to compactly express the difference in performance between two policies π' and π as

$$J_r(\pi') - J_r(\pi) = \frac{1}{1-\gamma} \mathbb{E}_{s \sim d^{\pi'}, a \sim \pi'}[A_r^\pi(s, a)],$$

where we use the shorthand $a \sim \pi'$ for $a \sim \pi'(\cdot|s)$. See Kakade and Langford (2002) for the proof of this identity.

Safe RL, crucial for safety-critical power system applications, extends standard RL by incorporating safety constraints, formalized through Constrained Markov Decision Processes (CMDPs) (Altman, 2021). A CMDP is represented as $\mathcal{M} \cup \mathcal{C}$, where $\mathcal{C} := (c, \xi)$ is the constraint tuple. Here, $c : \mathcal{S} \times \mathcal{A} \to \mathbb{R}$ denotes the cost function associated with safety violations, and ξ is the corresponding cost threshold. While we consider single cost function for simplicity of presentation, multiple constraints can be incorporated with individual cost function and threshold. We define on-policy value functions V_c^π, action-value functions Q_c^π, and advantage functions A_c^π for the cost in analogy to V_r^π, Q_r^π, and A_r^π with c replacing r in their respective definitions.

The safe RL objective is to find a policy $\pi : \mathcal{S} \to \Delta(\mathcal{A})$ that maximizes the expected cumulative reward $J_r(\pi)$ while adhering to the safety constraint:

$$\max_{\pi \in \Pi} J_r(\pi) \quad \text{subject to} \quad \pi \in \Pi_{\text{safe}} \tag{1.2}$$

where Π is the set of all policies. Various safety formulations of $\pi \in \Pi_{\text{safe}}$ can be considered:

1. Expected Cumulative Safety Constraint: $\mathbb{E}_\pi \left[\sum_{t=0}^{\infty} \gamma^t c(s_t, a_t)\right] \leq \xi$. This ensures that the expected cumulative cost remains below a threshold ξ; suitable for applications where occasional breaches are acceptable if the long-term average stays within safe limits, such as managing thermal loading, battery lifecycle, carbon emissions, or user comfort.

2. Expected Instantaneous Safety Constraint: $\mathbb{E}_\pi[c(s_t, a_t)] \leq \xi_t, \forall t$. This ensures the expected instantaneous cost remains below a threshold ξ_t at all times; suitable for applications where near-constant safety is crucial but occasional deviations are tolerable, such as managing voltage levels or EV charging rates.

3. Almost Surely Cumulative Safety Constraint: $\mathbb{P}_\pi[\sum_{t=0}^{\infty} \gamma^t c(s_t, a_t) \leq \xi] = 1$, where $\mathbb{P}_\pi(\cdot)$ represents probability under π and stochastic transition dynamics. This guarantees long-term safety with absolute certainty. It mandates that the cumulative cost remains

below a threshold ξ for *all possible trajectories* under policy π; essential for critical applications like ensuring trajectory-wise grid stability, where even rare violations can have severe consequences.

4. Almost Surely Instantaneous Safety Constraint: $\mathbb{P}_\pi[c(s_t, a_t) \leq \xi_t] = 1, \quad \forall t$. This is the strictest safety guarantee, demanding that the instantaneous cost remains below a threshold ξ_t with absolute certainty at every time step; crucial for critical safety parameters in power systems, such as maintaining grid frequency within strict limits or ensuring every action during critical load restoration is safe and avoids further system damage.

The State Constraint can be applied to any constraint type, where the cost function directly penalizing entry into unsafe states $c(s, a) = \mathbb{I}(s \in \mathcal{S}_{\text{unsafe}})$, where $\mathcal{S}_{\text{unsafe}} \subset S$ is the set of unsafe states and $\mathbb{I}(\cdot)$ is the indicator function.

Cumulative constraints (1, 3) prioritize long-term average performance, allowing for temporary violations if compensated over time. They are suitable for slow-changing processes and systems where operational flexibility is needed. Instantaneous constraints (2, 4), on the other hand, ensure safety at every time step, which is crucial for fast-dynamic systems where even brief violations are critical. The choice between these should be guided by the system's dynamics, the criticality of immediate safety, and the need for operational flexibility.

Expectation-based constraints (1, 2) offer more flexibility and are generally easier to implement and solve computationally. They allow for occasional violations, making them suitable for less critical parameters or systems with some tolerance for safety breaches. This approach often leads to policies with greater operational freedom and can be advantageous in multi-objective scenarios where strict safety might overly constrain other important objectives. In contrast, probability-based (Almost Surely) constraints (3, 4) provide stronger, trajectory-wise guarantees, ensuring no violations occur.[1] They are appropriate for critical safety parameters and align well with strict regulatory

[1] Probability-based constraints imply expectation-based ones: $\mathbb{P}_\pi[c(s_t, a_t) \leq \xi_t] = 1 \implies \mathbb{E}_\pi[c(s_t, a_t)] \leq \xi_t$. Conversely, expectation-based constraints can approximate probability-based ones: $\mathbb{E}_\pi[c(s_t, a_t)] \leq \frac{\xi_t}{\kappa} \implies \mathbb{P}_\pi[c(s_t, a_t) > \xi_t] \leq \frac{1}{\kappa}$ by Markov's inequality.

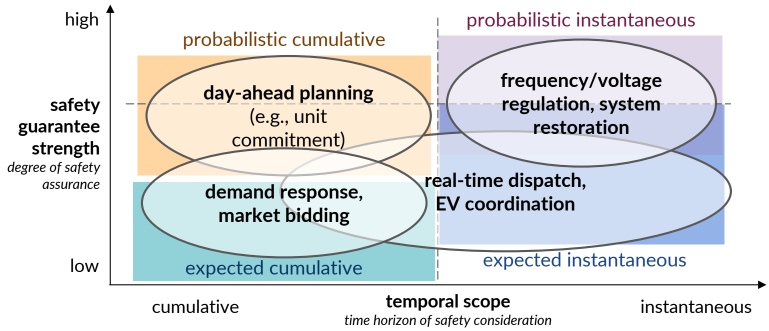

Figure 1.2: Safety constraint selection in power system applications spans a spectrum, from the most stringent for critical real-time operations like frequency/voltage regulation and system restoration, to more moderate levels for long-term planning and scheduling, and intermediate levels for grid-user interface management such as EV charging coordination. This adaptability of safe RL showcases its ability to balance the need for safety with the diverse operational requirements of power systems, ranging from strict real-time control to flexible long-term planning.

frameworks. However, these constraints may lead to more conservative policies and are typically more computationally intensive to implement and solve. The decision should consider the criticality of the safety parameter, regulatory requirements, available computational resources, and the system's tolerance for violations.

In power systems, these safety constraint formulations find application in a wide range of control and optimization problems, balancing efficiency and safety. Critical, fast-acting systems may require probability-based instantaneous constraints, while less critical, slower-changing aspects can utilize expectation-based cumulative constraints. The choice of formulation depends on factors such as safety requirements, system dynamics, computational resources, and uncertainty characterization, often benefiting from a combination of these approaches.

For example, Risk-Aware MDPs (RA-MDPs) introduce risk measures such as Conditional Value-at-Risk (CVaR) to model safety risk: $\mathrm{CVaR}_\beta\left(\sum_{t=0}^{\infty}\gamma^t c(s_t, a_t)\right) \leq \xi$. This constraint can be viewed as a variant of probability-based cumulative safety constraint and has been applied for managing risks associated with renewable energy integration and demand uncertainty (Yu *et al.*, 2024). Wu *et al.* (2024) apply probabilistic constraints to manage voltage levels, line thermal limits, and ensure grid stability under high DER penetration, addressing both instantaneous and dynamic violations. These approaches offer more flexible safety management, allowing for occasional constraint violations while maintaining probabilistic guarantees. This is crucial in power systems where strict constraints may lead to overly conservative or infeasible solutions, particularly in the presence of uncertainties from renewable sources and dynamic loads. Figure 1.2 provides a visual representation organized by temporal scope and safety guarantee strength.

References

Abdeen, Z. ul, X. Zhang, W. Gill, and M. Jin. (2024). "Enhancing Distribution System Resilience: A First-Order Meta-RL algorithm for Critical Load Restoration". In: *2024 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*. IEEE. 129–134.

Achiam, J., D. Held, A. Tamar, and P. Abbeel. (2017). "Constrained policy optimization". In: *International conference on machine learning*. PMLR. 22–31.

Ackermann, J., V. Gabler, T. Osa, and M. Sugiyama. (2019). "Reducing overestimation bias in multi-agent domains using double centralized critics". *arXiv preprint arXiv:1910.01465*.

Altman, E. (2021). *Constrained Markov decision processes*. Routledge.

Arcak, M., C. Meissen, and A. Packard. (2018). "Networks of Dissipative Systems".

Bai, N., Q. Wang, Z. Duan, and G. Chen. (2024). "Distributed Optimal Consensus Control of Constrained Multi-Agent Systems: A Non-Separable Optimization Perspective". *IEEE Transactions on Automatic Control*.

Bansal, S., M. Chen, S. Herbert, and C. J. Tomlin. (2017). "Hamilton-jacobi reachability: A brief overview and recent advances". In: *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*. IEEE. 2242–2253.

Beikbabaei, M., B. M. Kwiatkowski, and A. Mehrizi-Sani. (2025). "Model-Free Resilient Grid-Forming and Grid-Following Inverter Control Against Cyberattacks Using Reinforcement Learning". *Electronics*. 14(2): 288.

Biagioni, D., X. Zhang, D. Wald, D. Vaidhynathan, R. Chintala, J. King, and A. S. Zamzam. (2022). "Powergridworld: A framework for multi-agent reinforcement learning in power systems". In: *Proceedings of the thirteenth ACM international conference on future energy systems*. 565–570.

Biswas, A., M. Acquarone, H. Wang, F. Miretti, D. A. Misul, and A. Emadi. (2024). "Safe Reinforcement Learning for Energy Management of Electrified Vehicle with Novel Physics-Informed Exploration Strategy". *IEEE Transactions on Transportation Electrification*.

Bommasani, R., D. A. Hudson, E. Adeli, R. Altman, S. Arora, S. von Arx, M. S. Bernstein, J. Bohg, A. Bosselut, E. Brunskill, *et al.* (2021). "On the opportunities and risks of foundation models". *arXiv preprint arXiv:2108.07258*.

Brunke, L., M. Greeff, A. W. Hall, Z. Yuan, S. Zhou, J. Panerati, and A. P. Schoellig. (2022). "Safe learning in robotics: From learning-based control to safe reinforcement learning". *Annual Review of Control, Robotics, and Autonomous Systems*. 5(1): 411–444.

Chang, Y.-C., N. Roohi, and S. Gao. (2019). "Neural lyapunov control". *Advances in neural information processing systems*. 32.

Chen, D., K. Chen, Z. Li, T. Chu, R. Yao, F. Qiu, and K. Lin. (2021). "Powernet: Multi-agent deep reinforcement learning for scalable powergrid control". *IEEE Transactions on Power Systems*. 37(2): 1007–1017.

Chen, P., S. Liu, X. Wang, and I. Kamwa. (2022). "Physics-shielded multi-agent deep reinforcement learning for safe active voltage control with photovoltaic/battery energy storage systems". *IEEE Transactions on Smart Grid*. 14(4): 2656–2667.

Chen, Y., S. Huang, F. Liu, Z. Wang, and X. Sun. (2018). "Evaluation of reinforcement learning-based false data injection attack to automatic voltage control". *IEEE Transactions on Smart Grid*. 10(2): 2158–2169.

Chen, Y., Y. Liu, H. Yin, Z. Tang, G. Qiu, and J. Liu. (2024). "Multiagent Soft Actor–Critic Learning for Distributed ESS Enabled Robust Voltage Regulation of Active Distribution Grids". *IEEE Transactions on Industrial Informatics*.

Chen, Y. (2023). "High Performance Adaptive Deep-Reinforcement-Learning-based Real-time Emergency Control (HADREC) to Enhance Power Grid Resilience in Stochastic Environment Final Report". *Tech. rep.* Pacific Northwest National Laboratory (PNNL), Richland, WA (United States).

Cui, W., Y. Jiang, and B. Zhang. (2023). "Reinforcement learning for optimal primary frequency control: A Lyapunov approach". *IEEE Transactions on Power Systems*. 38(2): 1676–1688.

Du, Y. and D. Wu. (2022). "Deep reinforcement learning from demonstrations to assist service restoration in islanded microgrids". *IEEE Transactions on Sustainable Energy*. 13(2): 1062–1072.

Fan, B., X. Liu, G. Xiao, Y. Kang, D. Wang, and P. Wang. (2023a). "Attention-Based Multi-Agent Graph Reinforcement Learning for Service Restoration". *IEEE Transactions on Artificial Intelligence*.

Fan, B., X. Liu, G. Xiao, Y. Xu, X. Yang, and P. Wang. (2024). "A Memory-Based Graph Reinforcement Learning Method for Critical Load Restoration With Uncertainties of Distributed Energy Resource". *IEEE Transactions on Smart Grid*.

Fan, T.-H., X. Y. Lee, and Y. Wang. (2022). "Powergym: A reinforcement learning environment for volt-var control in power distribution systems". In: *Learning for Dynamics and Control Conference*. PMLR. 21–33.

Fan, Z., W. Zhang, and W. Liu. (2023b). "Multi-agent deep reinforcement learning based distributed optimal generation control of DC microgrids". *IEEE Transactions on Smart Grid*.

Feng, J., Y. Shi, G. Qu, S. H. Low, A. Anandkumar, and A. Wierman. (2023). "Stability constrained reinforcement learning for decentralized real-time voltage control". *IEEE Transactions on Control of Network Systems*.

Foerster, J., I. A. Assael, N. De Freitas, and S. Whiteson. (2016). "Learning to communicate with deep multi-agent reinforcement learning". *Advances in neural information processing systems*. 29.

Foerster, J., G. Farquhar, T. Afouras, N. Nardelli, and S. Whiteson. (2018). “Counterfactual multi-agent policy gradients”. In: *Proceedings of the AAAI conference on artificial intelligence*. Vol. 32. No. 1.

Gao, Y., W. Wang, and N. Yu. (2021). “Consensus multi-agent reinforcement learning for volt-var control in power distribution networks”. *IEEE Transactions on Smart Grid*. 12(4): 3594–3604.

Gao, Y. and N. Yu. (2022). “Model-augmented safe reinforcement learning for Volt-VAR control in power distribution networks”. *Applied Energy*. 313: 118762.

Gautam, M. (2023). “Deep Reinforcement learning for resilient power and energy systems: Progress, prospects, and future avenues”. *Electricity*. 4(4): 336–380.

Gu, F., H. Yin, L. El Ghaoui, M. Arcak, P. Seiler, and M. Jin. (2022). “Recurrent neural network controllers synthesis with stability guarantees for partially observed systems”. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 36. No. 5. 5385–5394.

Gu, S., J. G. Kuba, Y. Chen, Y. Du, L. Yang, A. Knoll, and Y. Yang. (2023). “Safe multi-agent reinforcement learning for multi-robot control”. *Artificial Intelligence*. 319: 103905.

Gu, S., B. Sel, Y. Ding, L. Wang, Q. Lin, M. Jin, and A. Knoll. (2024a). “Balance reward and safety optimization for safe reinforcement learning: A perspective of gradient manipulation”. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 38. No. 19. 21099–21106.

Gu, S., L. Shi, Y. Ding, A. Knoll, C. Spanos, A. Wierman, and M. Jin. (2024b). “Enhancing Efficiency of Safe Reinforcement Learning via Sample Manipulation”. *The Thirty-eighth Annual Conference on Neural Information Processing Systems (NeurIPS)*.

Guo, D., D. Yang, H. Zhang, J. Song, R. Zhang, R. Xu, Q. Zhu, S. Ma, P. Wang, X. Bi, *et al.* (2025). “Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning”. *arXiv preprint arXiv:2501.12948*.

Guo, G., M. Zhang, Y. Gong, and Q. Xu. (2023). “Safe multi-agent deep reinforcement learning for real-time decentralized control of inverter based renewable energy resources considering communication delay”. *Applied Energy*. 349: 121648.

Haarnoja, T., A. Zhou, P. Abbeel, and S. Levine. (2018). "Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor". In: *International conference on machine learning*. PMLR. 1861–1870.

Habib, B., E. Isufi, W. van Breda, A. Jongepier, and J. L. Cremer. (2023). "Deep statistical solver for distribution system state estimation". *IEEE Transactions on Power Systems*. 39(2): 4039–4050.

Hamann, H. F., B. Gjorgiev, T. Brunschwiler, L. S. Martins, A. Puech, A. Varbella, J. Weiss, J. Bernabe-Moreno, A. B. Massé, S. L. Choi, *et al.* (2024). "Foundation models for the electric power grid". *Joule*. 8(12): 3245–3258.

He, H., J. Boyd-Graber, K. Kwok, and H. Daumé III. (2016). "Opponent modeling in deep reinforcement learning". In: *International conference on machine learning*. PMLR. 1804–1813.

Henry, R. and D. Ernst. (2021). "Gym-ANM: Reinforcement learning environments for active network management tasks in electricity distribution systems". *Energy and AI*. 5: 100092.

Hernandez-Leal, P., M. Kaisers, T. Baarslag, and E. M. De Cote. (2017). "A survey of learning in multiagent environments: Dealing with non-stationarity". *arXiv preprint arXiv:1707.09183*.

Hewing, L., K. P. Wabersich, M. Menner, and M. N. Zeilinger. (2020). "Learning-based model predictive control: Toward safe learning in control". *Annual Review of Control, Robotics, and Autonomous Systems*. 3(1): 269–296.

Hobbs, K. L., M. L. Mote, M. C. Abate, S. D. Coogan, and E. M. Feron. (2023). "Runtime assurance for safety-critical systems: An introduction to safety filtering approaches for complex control systems". *IEEE Control Systems Magazine*. 43(2): 28–65.

Hong, L., M. Wu, Y. Wang, M. Shahidehpour, Z. Chen, and Z. Yan. (2024). "MADRL-Based DSO-Customer Coordinated Bi-Level Volt/-VAR Optimization Method for Power Distribution Networks". *IEEE Transactions on Sustainable Energy*.

Hong, S.-H. and H.-S. Lee. (2023). "Robust energy management system with safe reinforcement learning using short-horizon forecasts". *IEEE Transactions on Smart Grid*. 14(3): 2485–2488.

Hou, S., E. M. Salazar, P. Palensky, Q. Chen, and P. P. Vergara. (2024). "A mix-integer programming based deep reinforcement learning framework for optimal dispatch of energy storage system in distribution networks". *Journal of Modern Power Systems and Clean Energy*.

Hsu, K.-C., H. Hu, and J. F. Fisac. (2023). "The safety filter: A unified view of safety-critical control in autonomous systems". *Annual Review of Control, Robotics, and Autonomous Systems*. 7.

Hu, C., J. Yan, and X. Liu. (2022a). "Reinforcement learning-based adaptive feature boosting for smart grid intrusion detection". *IEEE Transactions on Smart Grid*. 14(4): 3150–3163.

Hu, D., Z. Ye, Y. Gao, Z. Ye, Y. Peng, and N. Yu. (2022b). "Multi-agent deep reinforcement learning for voltage control with coordinated active and reactive power optimization". *IEEE Transactions on Smart Grid*. 13(6): 4873–4886.

Hu, J., Y. Ye, Y. Wu, P. Zhao, and L. Liu. (2024). "Rethinking Safe Policy Learning for Complex Constraints Satisfaction: A Glimpse in Real-Time Security Constrained Economic Dispatch Integrating Energy Storage Units". *IEEE Transactions on Power Systems*.

Hu, L., C. Wu, and W. Pan. (2020). "Lyapunov-based reinforcement learning state estimator". *arXiv preprint arXiv:2010.13529*.

Iqbal, S. and F. Sha. (2019). "Actor-attention-critic for multi-agent reinforcement learning". In: *International conference on machine learning*. PMLR. 2961–2970.

Jacob, R. A., S. Paul, S. Chowdhury, Y. R. Gel, and J. Zhang. (2024). "Real-time outage management in active distribution networks using reinforcement learning over graphs". *Nature Communications*. 15(1): 4766.

Jeon, S., H. T. Nguyen, and D.-H. Choi. (2023). "Safety-Integrated Online Deep Reinforcement Learning for Mobile Energy Storage System Scheduling and Volt/VAR Control in Power Distribution Networks". *IEEE Access*. 11: 34440–34455.

Jiang, J. and Z. Lu. (2018). "Learning attentional communication for multi-agent cooperation". *Advances in neural information processing systems*. 31.

Jiang, Y., Q. Ye, B. Sun, Y. Wu, and D. H. Tsang. (2021). "Data-driven coordinated charging for electric vehicles with continuous charging rates: A deep policy gradient approach". *IEEE Internet of Things Journal*. 9(14): 12395–12412.

Jin, M. (2024). "Preparing for Black Swans: The Antifragility Imperative for Machine Learning". *arXiv preprint arXiv:2405.11397*.

Jin, M., V. Khattar, H. Kaushik, B. Sel, and R. Jia. (2023). "On solution functions of optimization: Universal approximation and covering number bounds". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 37. No. 7. 8123–8131.

Jin, M. and J. Lavaei. (2020). "Stability-certified reinforcement learning: A control-theoretic perspective". *IEEE Access*. 8: 229086–229100.

Jin, M., J. Lavaei, S. Sojoudi, and R. Baldick. (2020). "Boundary defense against cyber threat for power system state estimation". *IEEE Transactions on Information Forensics and Security*. 16: 1752–1767.

Jin, M., B. Sel, F. Hardeep, and W. Yin. (2024). "Democratizing energy management with llm-assisted optimization autoformalism". In: *2024 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*. IEEE. 258–263.

Kabir, F., N. Yu, Y. Gao, and W. Wang. (2023). "Deep reinforcement learning-based two-timescale Volt-VAR control with degradation-aware smart inverters in power distribution systems". *Applied Energy*. 335: 120629.

Kakade, S. and J. Langford. (2002). "Approximately optimal approximate reinforcement learning". In: *Proceedings of the Nineteenth International Conference on Machine Learning*. 267–274.

Kesici, M., B. Pal, and G. Yang. (2024). "Detection of false data injection attacks in distribution networks: A vertical federated learning approach". *IEEE Transactions on Smart Grid*.

Khattar, V., Y. Ding, B. Sel, J. Lavaei, and M. Jin. (2023). "A CMDP-within-online framework for Meta-Safe Reinforcement Learning". In: *The Eleventh International Conference on Learning Representations*.

Khattar, V. and M. Jin. (2023). "Winning the CityLearn challenge: adaptive optimization with evolutionary search under trajectory-based guidance". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 37. No. 12. 14286–14294.

Khattar, V., Y. Yao, F. Ding, and M. Jin. (2025). "Distribution Grid Critical Load Restoration under Uncertain Topology Changes via a Hierarchical Multi-Agent Reinforcement Learning Approach". In: *IEEE Power & Energy Society General Meeting*. IEEE. 1–5.

Kurt, M. N., O. Ogundijo, C. Li, and X. Wang. (2018). "Online cyber-attack detection in smart grid: A reinforcement learning approach". *IEEE Transactions on Smart Grid*. 10(5): 5174–5185.

Kwon, K.-B., R. R. Hossain, S. Mukherjee, K. Chatterjee, S. Kundu, S. Nekkalapu, and M. Elizondo. (2024). "Coherency-Aware Learning Control of Inverter-Dominated Grids: A Distributed Risk-Constrained Approach". *IEEE Control Systems Letters*.

Kwon, K.-b., S. Mukherjee, T. L. Vu, and H. Zhu. (2023). "Risk-Constrained Reinforcement Learning for Inverter-Dominated Power System Controls". *IEEE Control Systems Letters*.

Lee, Z. J., S. Sharma, D. Johansson, and S. H. Low. (2021). "ACN-Sim: An Open-Source Simulator for Data-Driven Electric Vehicle Charging Research". *IEEE Transactions on Smart Grid*. 12(6): 5113–5123.

Li, C. and Z. Wu. (2024). "F-DQN: an optimized DQN for decision-making of generator start-up sequence after blackout". *Applied Intelligence*: 1–15.

Li, H. and H. He. (2022). "Learning to operate distribution networks with safe deep reinforcement learning". *IEEE Transactions on Smart Grid*. 13(3): 1860–1872.

Li, H., Z. Wan, and H. He. (2019). "Constrained EV charging scheduling based on safe deep reinforcement learning". *IEEE Transactions on Smart Grid*. 11(3): 2427–2439.

Li, S., W. Hu, D. Cao, Z. Chen, Q. Huang, F. Blaabjerg, and K. Liao. (2023). "Physics-model-free heat-electricity energy management of multiple microgrids based on surrogate model-enabled multi-agent deep reinforcement learning". *Applied Energy*. 346: 121359.

Li, S., W. Hu, D. Cao, J. Hu, Q. Huang, Z. Chen, and F. Blaabjerg. (2024). "A Novel MADRL with Spatial-Temporal Pattern Capturing Ability for Robust Decentralized Control of Multiple Microgrids under Anomalous Measurements". *IEEE Transactions on Sustainable Energy*.

Lin, Y., W. Li, H. Zha, and B. Wang. (2024). "Information Design in Multi-Agent Reinforcement Learning". *Advances in Neural Information Processing Systems*. 36.

Liu, H. and W. Wu. (2021). "Online multi-agent reinforcement learning for decentralized inverter-based volt-var control". *IEEE Transactions on Smart Grid*. 12(4): 2980–2990.

Liu, L., N. Shi, D. Wang, Z. Ma, Z. Wang, M. J. Reno, and J. A. Azzolini. (2024a). "Voltage calculations in secondary distribution networks via physics-inspired neural network using smart meter data". *IEEE Transactions on Smart Grid*.

Liu, X., Q. Jiao, S. Qiao, Z. Yan, S. Wen, and P. Wang. (2024b). "A Hybrid Monotonic Neural Network Approach for Multi-Area Power System Load Frequency Control Against FGSM Attack". *IEEE Transactions on Circuits and Systems II: Express Briefs*.

Lohmiller, W. and J.-J. E. Slotine. (1998). "On contraction analysis for non-linear systems". *Automatica*. 34(6): 683–696.

Lowe, R., Y. I. Wu, A. Tamar, J. Harb, O. Pieter Abbeel, and I. Mordatch. (2017). "Multi-agent actor-critic for mixed cooperative-competitive environments". *Advances in neural information processing systems*. 30.

Luo, S., M. Jin, R. Han, and K. Sun. (2022). "Anti-jamming based on Reinforcement Learning in Power System Sensing Network". In: *Proceedings of the 2022 12th International Conference on Communication and Network Security*. 180–184.

Ma, D., Y. Wang, and S. Wu. (2024). "Against Jamming Attack in Wireless Communication Networks: A Reinforcement Learning Approach". *Electronics*. 13(7): 1209.

Maiti, S. and S. Dey. (2024). "Smart Grid Security: A Verified Deep Reinforcement Learning Framework to Counter Cyber-Physical Attacks". *arXiv preprint arXiv:2409.15757*.

Megretski, A. and A. Rantzer. (1997). "System analysis via integral quadratic constraints". *IEEE transactions on automatic control*. 42(6): 819–830.

Moradi, M., S. Panahi, Z.-M. Zhai, Y. Weng, J. Dirkman, and Y.-C. Lai. (2024). "Heterogeneous reinforcement learning for defending power grids against attacks". *APL Machine Learning*. 2(2).

Moradi, M., Y. Weng, J. Dirkman, and Y.-C. Lai. (2023). "Preferential cyber defense for power grids". *PRX Energy*. 2(4): 043007.

Mouyart, M., G. Medeiros Machado, and J.-Y. Jun. (2023). "A multi-agent intrusion detection system optimized by a deep reinforcement learning approach with a dataset enlarged using a generative model to reduce the bias effect". *Journal of Sensor and Actuator Networks*. 12(5): 68.

Mu, C., Z. Liu, J. Yan, H. Jia, and X. Zhang. (2023). "Graph multi-agent reinforcement learning for inverter-based active voltage control". *IEEE Transactions on Smart Grid*.

Mukherjee, S., R. R. Hossain, S. M. Mohiuddin, Y. Liu, W. Du, V. Adetola, R. A. Jinsiwale, Q. Huang, T. Yin, and A. Singhal. (2024). "Resilient Control of Networked Microgrids using Vertical Federated Reinforcement Learning: Designs and Real-Time Test-Bed Validations". *IEEE Transactions on Smart Grid*.

Nagy, Z., J. R. Vázquez-Canteli, S. Dey, and G. Henze. (2021). "The citylearn challenge 2021". In: *Proceedings of the 8th ACM international conference on systems for energy-efficient buildings, cities, and transportation*. 218–219.

Nguyen, H. T. and D.-H. Choi. (2022). "Three-stage inverter-based peak shaving and Volt-VAR control in active distribution networks using online safe deep reinforcement learning". *IEEE Transactions on Smart Grid*. 13(4): 3266–3277.

Oliehoek, F. A. and C. Amato. (2016). *A concise introduction to decentralized POMDPs*. Vol. 1. Springer.

Omidshafiei, S., J. Pazis, C. Amato, J. P. How, and J. Vian. (2017). "Deep decentralized multi-task multi-agent reinforcement learning under partial observability". In: *International Conference on Machine Learning*. PMLR. 2681–2690.

Orfanoudakis, S., C. Diaz-Londono, Y. E. Yılmaz, P. Palensky, and P. P. Vergara. (2024). "Ev2gym: A flexible v2g simulator for ev smart charging research and benchmarking". *IEEE Transactions on Intelligent Transportation Systems*.

Paesschesoone, S., N. Kayedpour, C. Manna, and G. Crevecoeur. (2024). "Reinforcement learning for an enhanced energy flexibility controller incorporating predictive safety filter and adaptive policy updates". *Applied Energy*. 368: 123507.

Pigott, A., C. Crozier, K. Baker, and Z. Nagy. (2022). "Gridlearn: Multiagent reinforcement learning for grid-aware building energy management". *Electric power systems research*. 213: 108521.

Qin, Z., K. Zhang, Y. Chen, J. Chen, and C. Fan. (2021). "Learning safe multi-agent control with decentralized neural barrier certificates". *arXiv preprint arXiv:2101.05436*.

Qiu, D., Z. Dong, X. Zhang, Y. Wang, and G. Strbac. (2022). "Safe reinforcement learning for real-time automatic control in a smart energy-hub". *Applied Energy*. 309: 118403.

Qiu, D., Y. Wang, J. Wang, N. Zhang, G. Strbac, and C. Kang. (2023). "Resilience-Oriented Coordination of Networked Microgrids: A Shapley Q-Value Learning Approach". *IEEE transactions on power systems*. 39(2): 3401–3416.

Rashid, T., M. Samvelyan, C. S. De Witt, G. Farquhar, J. Foerster, and S. Whiteson. (2020). "Monotonic value function factorisation for deep multi-agent reinforcement learning". *Journal of Machine Learning Research*. 21(178): 1–51.

Ray, A., J. Achiam, and D. Amodei. (2019). "Benchmarking safe exploration in deep reinforcement learning". *arXiv preprint arXiv:1910.01708*. 7(1): 2.

Sahu, A., V. Venkatraman, and R. Macwan. (2023). "Reinforcement learning environment for cyber-resilient power distribution system". *IEEE Access*. 11: 127216–127228.

Salamat, B., G. Elsbacher, A. M. Tonello, and L. Belzner. (2023). "Model-free distributed reinforcement learning state estimation of a dynamical system using integral value functions". *IEEE Open Journal of Control Systems*. 2: 70–78.

Schierman, J. D., M. D. DeVore, N. D. Richards, N. Gandhi, J. K. Cooper, K. R. Horneman, S. Stoller, and S. Smolka. (2015). "Runtime assurance framework development for highly adaptive flight control systems". *Barron Associates, Inc. Charlottesville, Tech. Rep.*

Seiler, P. (2014). "Stability analysis with dissipation inequalities and integral quadratic constraints". *IEEE Transactions on Automatic Control*. 60(6): 1704–1709.

Sethi, K., Y. V. Madhav, R. Kumar, and P. Bera. (2021). "Attention based multi-agent intrusion detection systems using reinforcement learning". *Journal of Information Security and Applications*. 61: 102923.

Seto, D., B. Krogh, L. Sha, and A. Chutinan. (1998). "The simplex architecture for safe online control system upgrades". In: *Proceedings of the 1998 American Control Conference*. Vol. 6. IEEE. 3504–3508.

Shengren, H., P. P. Vergara, E. M. S. Duque, and P. Palensky. (2023). "Optimal energy system scheduling using a constraint-aware reinforcement learning algorithm". *International Journal of Electrical Power & Energy Systems*. 152: 109230.

Shi, Y., M. Feng, M. Wang, W. Zhou, and H. Li. (2023). "Multi-Agent Reinforcement Learning with Safety Layer for Active Voltage Control". In: *Proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems*. 1533–1541.

Shuai, H., B. She, J. Wang, and F. Li. (2024). "Safe Reinforcement Learning for Grid-Forming Inverter Based Frequency Regulation with Stability Guarantee". *Journal of Modern Power Systems and Clean Energy*.

Si, R., S. Chen, J. Zhang, J. Xu, and L. Zhang. (2024). "A multi-agent reinforcement learning method for distribution system restoration considering dynamic network reconfiguration". *Applied Energy*. 372: 123625.

Sootla, A., A. I. Cowen-Rivers, T. Jafferjee, Z. Wang, D. H. Mguni, J. Wang, and H. Ammar. (2022). "Sauté rl: Almost surely safe reinforcement learning using state augmentation". In: *International Conference on Machine Learning*. PMLR. 20423–20443.

Sukhbaatar, S. and R. Fergus. (2016). "Learning multiagent communication with backpropagation". *Advances in neural information processing systems*. 29.

Sun, X., Z. Xu, J. Qiu, H. Liu, H. Wu, and Y. Tao. (2024). "Optimal volt/var control for unbalanced distribution networks with human-in-the-loop deep reinforcement learning". *IEEE Transactions on Smart Grid*.

Sun, Z. and T. Lu. (2024). "Collaborative operation optimization of distribution system and virtual power plants using multi-agent deep reinforcement learning with parameter-sharing mechanism". *IET Generation, Transmission & Distribution*. 18(1): 39–49.

Sunehag, P., G. Lever, A. Gruslys, W. M. Czarnecki, V. Zambaldi, M. Jaderberg, M. Lanctot, N. Sonnerat, J. Z. Leibo, K. Tuyls, *et al.* (2018). "Value-Decomposition Networks For Cooperative Multi-Agent Learning Based On Team Reward". In: *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems*. 2085–2087.

Taleb, N. N. (2014). *Antifragile: Things that gain from disorder*. Vol. 3. Random House Trade Paperbacks.

Tan, X. and D. V. Dimarogonas. (2021). "Distributed implementation of control barrier functions for multi-agent systems". *IEEE Control Systems Letters*. 6: 1879–1884.

Tan, X., C. Liu, K. H. Johansson, and D. V. Dimarogonas. (2024). "A continuous-time violation-free multi-agent optimization algorithm and its applications to safe distributed control". *arXiv preprint arXiv:2404.07571*.

Tu, S., Y. Zhang, J. Zhang, Z. Fu, Y. Zhang, and Y. Yang. (2024). "Powerpm: Foundation model for power systems". *Advances in Neural Information Processing Systems*. 37: 115233–115260.

Vazquez-Canteli, J. R., S. Dey, G. Henze, and Z. Nagy. (2020). "CityLearn: Standardizing research in multi-agent reinforcement learning for demand response and urban energy management". *arXiv preprint arXiv:2012.10504*.

Vu, L., T. Vu, T. L. Vu, and A. Srivastava. (2023). "Multi-agent deep reinforcement learning for distributed load restoration". *IEEE Transactions on Smart Grid*.

Wan, X., M. Sun, B. Chen, Z. Chu, and F. Teng. (2023). "AdapSafe: adaptive and safe-certified deep reinforcement learning-based frequency control for carbon-neutral power systems". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 37. No. 4. 5294–5302.

Wang, J., Y. Zhang, T.-K. Kim, and Y. Gu. (2020). "Shapley Q-value: A local reward approach to solve global reward games". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 34. No. 05. 7285–7292.

Wang, L., S. Zhang, Y. Zhou, C. Fan, P. Zhang, and Y. A. Shamash. (2023a). "Physics-informed, safety and stability certified neural control for uncertain networked microgrids". *IEEE Transactions on Smart Grid*.

Wang, W., N. Yu, Y. Gao, and J. Shi. (2019). "Safe off-policy deep reinforcement learning algorithm for volt-var control in power distribution systems". *IEEE Transactions on Smart Grid*. 11(4): 3008–3018.

Wang, Y., D. Qiu, M. Sun, G. Strbac, and Z. Gao. (2023b). "Secure energy management of multi-energy microgrid: A physical-informed safe reinforcement learning approach". *Applied Energy*. 335: 120759.

Wang, Y., D. Qiu, X. Sun, Z. Bie, and G. Strbac. (2023c). "Coordinating multi-energy microgrids for integrated energy system resilience: A multi-task learning approach". *IEEE Transactions on Sustainable Energy*.

Wang, Y., Z. Yan, L. Sang, L. Hong, Q. Hu, M. Shahidehpour, and Q. Xu. (2023d). "Acceleration Framework and Solution Algorithm for Distribution System Restoration Based on End-to-End Optimization Strategy". *IEEE Transactions on Power Systems*. 39(1): 429–441.

Willems, J. C. (1972). "Dissipative dynamical systems part I: General theory". *Archive for rational mechanics and analysis*. 45(5): 321–351.

Williams, R. J. (1992). "Simple statistical gradient-following algorithms for connectionist reinforcement learning". *Machine learning*. 8: 229–256.

Wu, X., S. Magnússon, and M. Johansson. (2023). "Distributed safe resource allocation using barrier functions". *Automatica*. 153: 111051.

Wu, Y., Y. Ye, J. Hu, P. Zhao, L. Liu, G. Strbac, and C. Kang. (2024). "Chance Constrained MDP Formulation and Bayesian Advantage Policy Optimization for Stochastic Dynamic Optimal Power Flow". *IEEE Transactions on Power Systems*.

Xia, Y., Y. Xu, Y. Wang, S. Mondal, S. Dasgupta, A. K. Gupta, and G. M. Gupta. (2022). "A safe policy learning-based method for decentralized and economic frequency control in isolated networked-microgrid systems". *IEEE Transactions on Sustainable Energy*. 13(4): 1982–1993.

Xu, T., Y. Liang, and G. Lan. (2021). "Crpo: A new approach for safe reinforcement learning with convergence guarantee". In: *International Conference on Machine Learning*. PMLR. 11480–11491.

Xu, X., H. Hu, Y. Liu, J. Tan, H. Zhang, and H. Song. (2022). "Moving target defense of routing randomization with deep reinforcement learning against eavesdropping attack". *Digital Communications and Networks*. 8(3): 373–387.

Yan, R., Q. Xing, and Y. Xu. (2023). "Multi-Agent Safe Graph Reinforcement Learning for PV Inverters-Based Real-Time Decentralized Volt/Var Control in Zoned Distribution Networks". *IEEE Transactions on Smart Grid*. 15(1): 299–311.

Yan, Z. and Y. Xu. (2022). "A hybrid data-driven method for fast solution of security-constrained optimal power flow". *IEEE Transactions on Power Systems*. 37(6): 4365–4374.

Yang, L., G. Chen, and X. Cao. (2025). "A deep reinforcement learning-based charging scheduling approach with augmented Lagrangian for electric vehicles". *Applied Energy*. 378: 124706.

Yang, T. Y., J. Rosca, K. Narasimhan, and P. J. Ramadge. (2020). "Projection-Based Constrained Policy Optimization". In: *8th International Conference on Learning Representations (ICLR)*.

Yeh, C., V. Li, R. Datta, J. Arroyo, N. Christianson, C. Zhang, Y. Chen, M. M. Hosseini, A. Golmohammadi, Y. Shi, *et al.* (2024). "SustainGym: Reinforcement learning environments for sustainable energy systems". *Advances in Neural Information Processing Systems*. 36.

Ying, D., Y. Zhang, Y. Ding, A. Koppel, and J. Lavaei. (2024). "Scalable primal-dual actor-critic method for safe multi-agent rl with general utilities". *Advances in Neural Information Processing Systems*. 36.

Yu, P., H. Zhang, and Y. Song. (2024). "Adaptive Tie-Line Power Smoothing With Renewable Generation Based on Risk-Aware Reinforcement Learning". *IEEE Transactions on Power Systems*.

Yu, P., H. Zhang, Y. Song, H. Hui, and G. Chen. (2023). "District cooling system control for providing operating reserve based on safe deep reinforcement learning". *IEEE Transactions on Power Systems*. 39(1): 40–52.

Yuan, Y., K. Dehghanpour, Z. Wang, and F. Bu. (2022). "A joint distribution system state estimation framework via deep actor-critic learning method". *IEEE Transactions on Power Systems*. 38(1): 796–806.

Yuan, Z., C. Zhao, and J. Cortés. (2024). "Reinforcement learning for distributed transient frequency control with stability and safety guarantees". *Systems & Control Letters*. 185: 105753.

Zhang, H., J. Peng, H. Tan, H. Dong, and F. Ding. (2020a). "A deep reinforcement learning-based energy management framework with Lagrangian relaxation for plug-in hybrid electric vehicle". *IEEE Transactions on Transportation Electrification*. 7(3): 1146–1160.

Zhang, J., L. Sang, Y. Xu, and H. Sun. (2024a). "Networked Multiagent-Based Safe Reinforcement Learning for Low-Carbon Demand Management in Distribution Networks". *IEEE Transactions on Sustainable Energy*.

Zhang, J., Y. Guan, L. Che, and M. Shahidehpour. (2023a). "EV charging command fast allocation approach based on deep reinforcement learning with safety modules". *IEEE Transactions on Smart Grid*. 15(1): 757–769.

Zhang, M., G. Guo, S. Magnússon, R. C. Pilawa-Podgurski, and Q. Xu. (2023b). "Data driven decentralized control of inverter based renewable energy sources using safe guaranteed multi-agent deep reinforcement learning". *IEEE Transactions on Sustainable Energy*.

Zhang, M., G. Guo, T. Zhao, and Q. Xu. (2023c). "Dnn assisted projection based deep reinforcement learning for safe control of distribution grids". *IEEE Transactions on Power Systems*.

Zhang, Q., K. Dehghanpour, Z. Wang, F. Qiu, and D. Zhao. (2020b). "Multi-agent safe policy learning for power management of networked microgrids". *IEEE Transactions on Smart Grid*. 12(2): 1048–1062.

Zhang, S., R. Jia, H. Pan, and Y. Cao. (2023d). "A safe reinforcement learning-based charging strategy for electric vehicles in residential microgrid". *Applied Energy*. 348: 121490.

Zhang, S., K. Garg, and C. Fan. (2023e). "Neural graph control barrier functions guided distributed collision-avoidance multi-agent control". In: *Conference on robot learning*. PMLR. 2373–2392.

Zhang, X., Q. Wang, X. Bi, D. Li, D. Liu, Y. Yu, and C. K. Tse. (2024b). "Mitigating cascading failure in power grids with deep reinforcement learning-based remedial actions". *Reliability Engineering & System Safety*. 250: 110242.

Zhang, Y., M. Li, Y. Chen, Y. Y. Chiang, and Y. Hua. (2023f). "A Constraint-Based Routing and Charging Methodology for Battery Electric Vehicles With Deep Reinforcement Learning". *IEEE Transactions on Smart Grid*. 14(3): 2446–2459.

Zhang, Y., J. Zhao, D. Shi, and S. Chung. (2024c). "Deep Reinforcement Learning-Enabled Adaptive Forecasting-Aided State Estimation in Distribution Systems with Multi-Source Multi-Rate Data". In: *2024 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*. IEEE. 1–5.

Zhang, Y., J. Hu, G. Min, X. Chen, and N. Georgalas. (2024d). "Joint Charging Scheduling and Computation Offloading in EV-Assisted Edge Computing: A Safe DRL Approach". *IEEE Transactions on Mobile Computing*.

Zhao, T., J. Wang, and M. Yue. (2023). "A barrier-certificated reinforcement learning approach for enhancing power system transient stability". *IEEE Transactions on Power Systems*. 38(6): 5356–5366.

Zhou, K. and J. C. Doyle. (1998). *Essentials of robust control*. Vol. 104. Prentice hall Upper Saddle River, NJ.