Deep Reinforcement Learning based Optimal Energy System Scheduling

Hou Shengren

TUDelft

Content

1. Self-Introduction
2. Introduction for RL and energy management
3. Safety problem induced by RL
4. Our motivation and experiments
5. My PhD Research Routine

Self-Introduction

1. Researcher (Power and financial market, Power System, AI)

2. Career (Quantitative Power Trader, Power market expert)

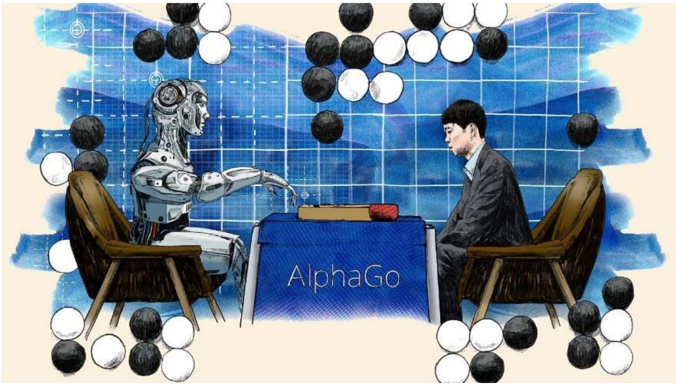
3. Entrepreneur (Energy Quant Research Institution)

4. Social activity (Board member of VCWI)

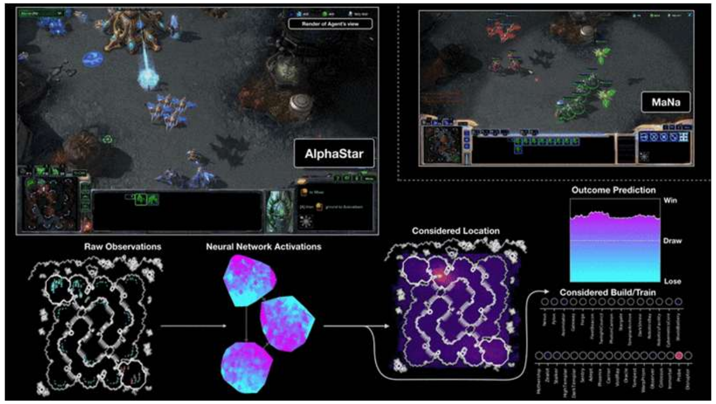
TU Delft IEPG Intelligent Electrical Power Grids

ENERGY QUANT 能源量化 Research Institution

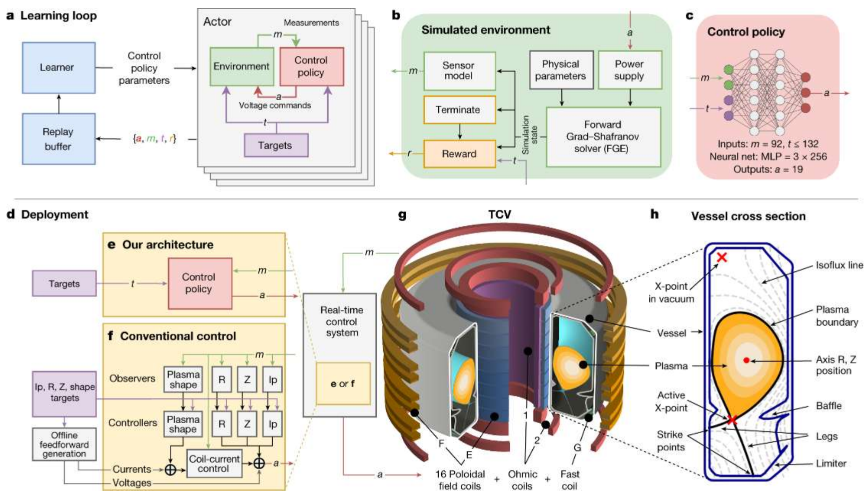
Reinforcement Learning Introduction



2014



2018

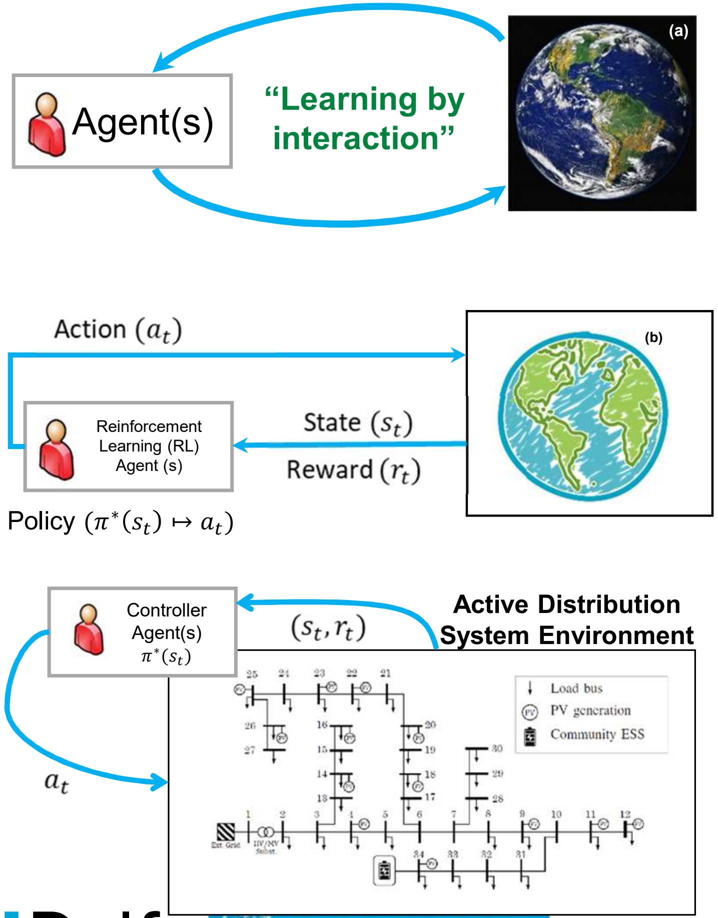


2022

TU Delft

IEPG Intelligent Electrical Power Grids

Active distribution networks (ADNs) operation as a RL problem



RL solves a sequential problem that is formulated as a Markov Decision Problem (MDP):

$$< S, A, P, r, \gamma > \quad \begin{cases} S\text{: State space (Observed variables)} \\ A\text{: Action space (possible control actions)} \\ P\text{: Transition probability (Not available but simulated)} \\ r\text{: Reward function (Signal to maximize)} \\ \gamma\text{: Discount factor (importance of future rewards)} \end{cases}$$

Advantages of RL:

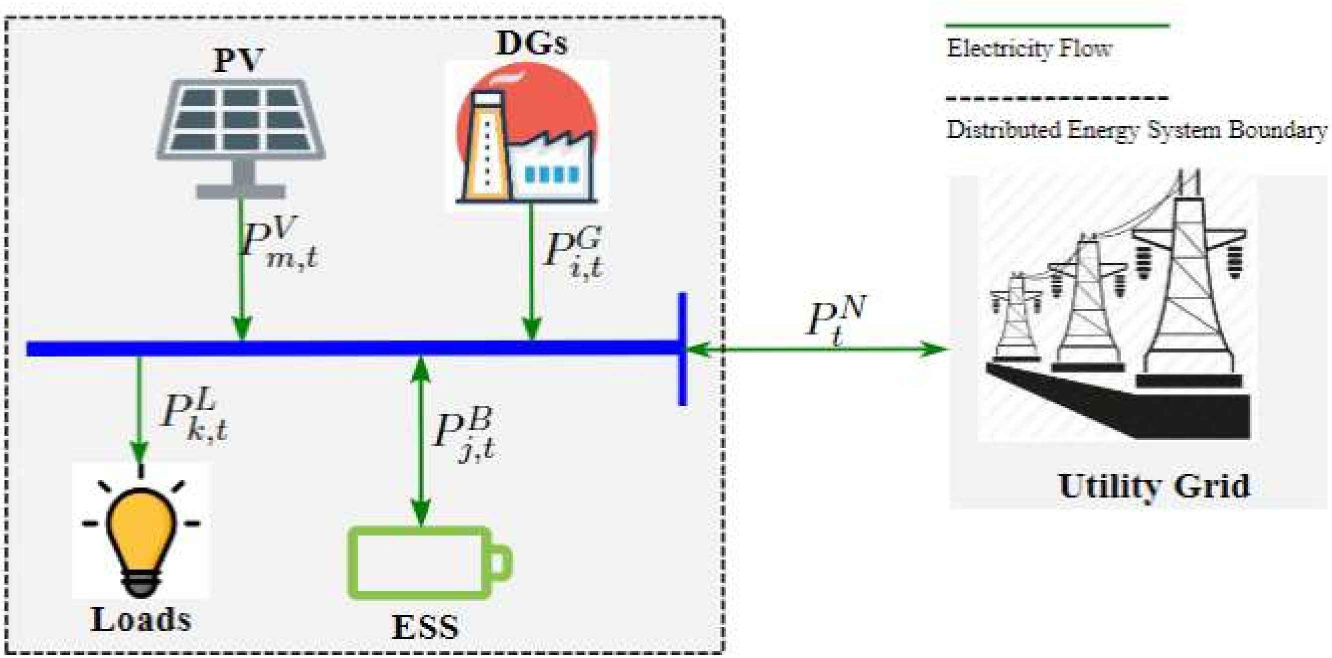
- It is a "model-free" approach to solve decision-making problems.
- Excellent generalization features. Optimal actions for different states.
- Complexity of the system (environment) can be high. Using powerful parametrized function approximators for $\pi(s, \theta)$ (e.g. Deep Neural Networks), we can find good and practical solutions.

TU Delft

IEPG Intelligent Electrical Power Grids

ENERGY QUANT 能源量化 Research Institution

Experiment Design



Research Background: What is optimal energy system scheduling?

$$\min \sum_{t \in \mathcal{T}} \sum_{i \in \mathcal{G}} (C_{i,t}^{G} + C_{t}^{E}) \Delta t \quad (1)$$

$$C_{i,t}^{G} = a_i \cdot (P_{i,t}^{G})^2 + b_i \cdot P_{i,t}^{G} + c_i, \quad i \in \mathcal{G}. \quad (2)$$

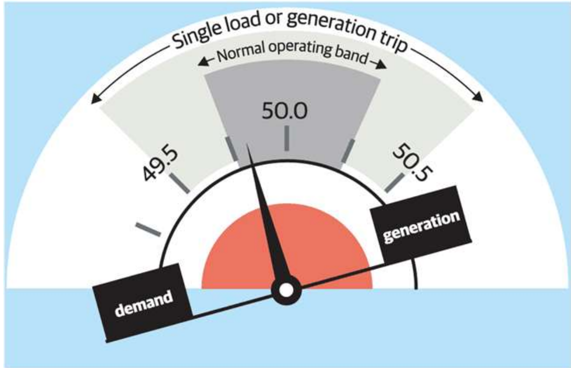
$$C_t^E = \begin{cases} \rho_t P_t^N & P_t^N > 0, \\ \beta \rho_t P_t^N & P_t^N < 0. \end{cases} \quad (3)$$

Minimize energy costs

Subject to:

$$\sum_{i \in \mathcal{G}} P_{i,t}^{G} + \sum_{m \in \mathcal{V}} P_{m,t}^{V} + P_t^N + \sum_{j \in \mathcal{B}} P_{j,t}^{B} = \sum_{k \in \mathcal{L}} P_{k,t}^{L}, \forall t \in \mathcal{T} \quad (4)$$

Power balance constrain



$$\underline{P}_i^G \cdot u_{i,t} \leq P_{i,t}^{G} \leq \overline{P}_i^G \cdot u_{i,t} \quad \forall i \in \mathcal{G}, \forall t \in \mathcal{T} \quad (5)$$

$$P_{i,t}^{G} - P_{i,t-1}^{G} \leq RU_i \quad \forall i \in \mathcal{G}, \forall t \in \mathcal{T} \quad (6)$$

$$P_{i,t}^{G} - P_{i,t+1}^{G} \leq RD_i, \quad \forall i \in \mathcal{G}, \forall t \in \mathcal{T} \quad (7)$$

$$-\underline{P}_j^B \leq P_{j,t}^{B} \leq \overline{P}_j^B \quad \forall j \in \mathcal{B}, \forall t \in \mathcal{T} \quad (8)$$

Technical limits for generators

Mathematical essence is to search for optimal solution for sequential decision problems within limited time window.

TU Delft IEPG Intelligent Electrical Power Grids

ENERGY QUANT 能源量化 Research Institution

Case Study: Energy System Optimal Scheduling

$$\min_{P_{i,t}^{G}, P_{j,t}^{B}} \left\{ \sum_{t\in\mathcal{T}} \sum_{i\in\mathcal{G}} \left[C_{i,t}^{G}(\cdot) + C_{t}^{E}(\cdot) \right] \Delta t \right\}, \quad (1)$$

$$C_{i,t}^{G} = a_i \left(P_{i,t}^{G}\right)^2 + b_i P_{i,t}^{G} + c_i, \quad \forall i \in \mathcal{G}. \quad (2)$$

$$C_t^E = \begin{cases} \rho_t P_t^N & P_t^N > 0, \\ \beta \rho_t P_t^N & P_t^N < 0. \end{cases} \quad (3)$$

Minimization of all operational cost

Subject to:

$$\sum_{i\in\mathcal{G}} P_{i,t}^{G} + \sum_{m\in\mathcal{V}} P_{m,t}^{V} + P_t^N + \sum_{j\in\mathcal{B}} P_{j,t}^{B} = \sum_{k\in\mathcal{L}} P_{k,t}^{L}, \forall t \in \mathcal{T} \quad (4)$$

Power balance constraint

$$\underline{P}_i^G \leq P_{i,t}^G \leq \overline{P}_i^G \quad \forall i \in \mathcal{G}, \forall t \in \mathcal{T} \quad (5)$$

$$P_{i,t}^G - P_{i,t-1}^G \leq RU_i \quad \forall i \in \mathcal{G}, \forall t \in \mathcal{T} \quad (6)$$

$$P_{i,t}^G - P_{i,t+1}^G \leq RD_i \quad \forall i \in \mathcal{G}, \forall t \in \mathcal{T} \quad (7)$$

$$-\underline{P}_j^B \leq P_{j,t}^B \leq \overline{P}_j^B \quad \forall j \in \mathcal{B}, \forall t \in \mathcal{T} \quad (8)$$

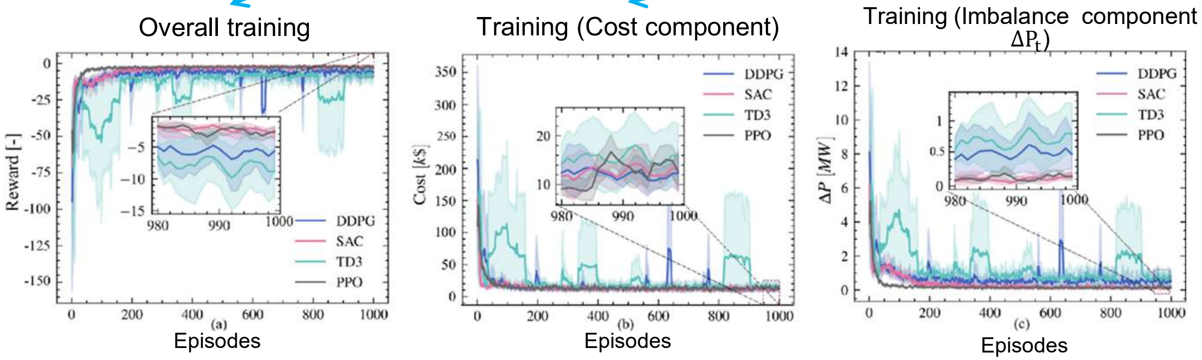
$$SOC_{j,t}^B = SOC_{j,t-1}^B + \eta_B P_{j,t}^B \Delta t / E_j^B \quad \forall j \in \mathcal{B}, \forall t \in \mathcal{T} \quad (9)$$

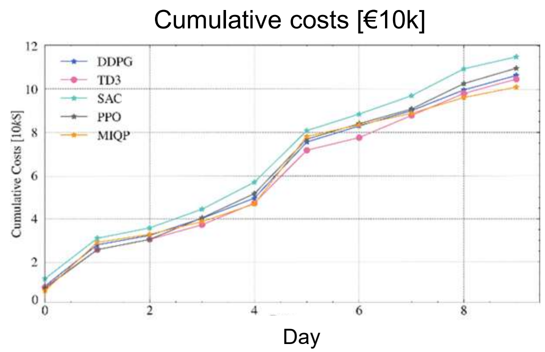
$$\underline{SOC}_j^B \leq SOC_{j,t}^B \leq \overline{SOC}_j^B \quad \forall j \in \mathcal{B}, \forall t \in \mathcal{T} \quad (10)$$

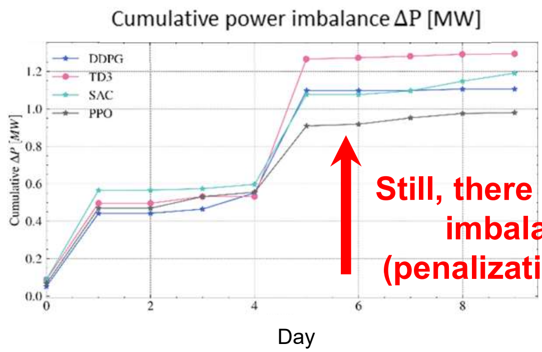
$$-\overline{P}^C \leq P_t^N \leq \overline{P}^C \quad \forall t \in \mathcal{T} \quad (11)$$

Operational constraints

$$r_t(s_t, a_t) = -\sigma_1 \left[\sum_{i\in\mathcal{G}} C_{i,t}^G + C_t^E \right] - \sigma_2 \Delta P_t$$







Still, there is power imbalance (penalization only)

(D)RL algorithms *lack of safety guarantees*, as they cannot (yet) be directly imposed in the algorithm's formulation.

TU Delft IEPG Intelligent Electrical Power Grids

"Performance Comparison of Deep RL Algorithms for Energy Systems Optimal Scheduling," Hou Shengren, Edgar Mauricio Sal Pedro P. Vergara, Peter Palensky, ISGT Europe 2022.

Our idea and experiments validation

Goal

Reinforcement learning algorithms that can provide theoretical proof of the constraint handling, during the energy management operation.

Background

- providing such proof of RL algorithms can be difficult and may not always feasible.
- This is because the lack of mathematical tools and theories for RL algorithms

Motivation

- Classic model based approaches like MPC, MILP have well-established mathematics theories. We can formulate the trained RL algorithm as a MIP, which can bring a stronger theoretical foundation for RL algorithms.
- In this way, Various mathematical theories can be used for ensuring the feasibility of our algorithm, such as duality theory, convex optimization, or polyhedral theory.

TU Delft IEPG Intelligent Electrical Power Grids

ENERGY QUANT 能源量化 Research Institution

Case Study: Energy System Optimal Scheduling

Understanding the (operational) constraints in the action space:

Subject to:

$$\sum_{i\in\mathcal{G}} P_{i,t}^{G} + \sum_{m\in\mathcal{V}} P_{m,t}^{V} + P_t^{N} + \sum_{j\in\mathcal{B}} P_{j,t}^{B} = \sum_{k\in\mathcal{L}} P_{k,t}^{L}, \forall t \in \mathcal{T} \tag{4}$$

$$\underline{P}_i^{G} \le P_{i,t}^{G} \le \overline{P}_i^{G} \qquad \forall i \in \mathcal{G}, \forall t \in \mathcal{T} \tag{5}$$

$$P_{i,t}^{G} - P_{i,t-1}^{G} \le RU_i \qquad \forall i \in \mathcal{G}, \forall t \in \mathcal{T} \tag{6}$$

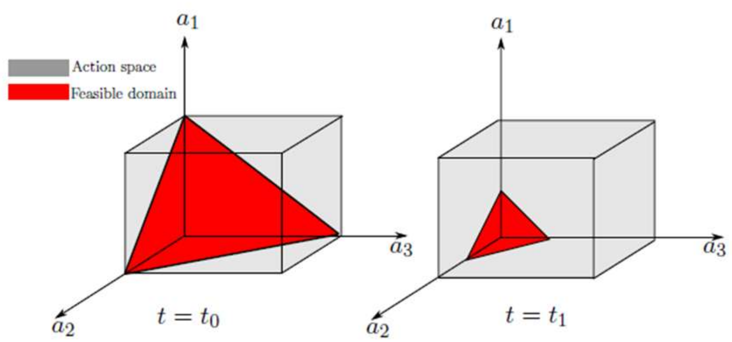
$$P_{i,t}^{G} - P_{i,t+1}^{G} \le RD_i \qquad \forall i \in \mathcal{G}, \forall t \in \mathcal{T} \tag{7}$$

$$-\underline{P}_j^{B} \le P_{j,t}^{B} \le \overline{P}_j^{B} \qquad \forall j \in \mathcal{B}, \forall t \in \mathcal{T} \tag{8}$$

$$SOC_{j,t}^{B} = SOC_{j,t-1}^{B} + \eta_B P_{j,t}^{B} \Delta t / E_j^{B} \qquad \forall j \in \mathcal{B}, \forall t \in \mathcal{T} \tag{9}$$

$$\underline{SOC}_j^{B} \le SOC_{j,t}^{B} \le \overline{SOC}_j^{B} \qquad \forall j \in \mathcal{B}, \forall t \in \mathcal{T} \tag{10}$$

$$-\overline{P}^{C} \le P_t^{N} \le \overline{P}^{C} \qquad \forall t \in \mathcal{T} \tag{11}$$



The equality constraint defines the feasible action space (red space, hyperplane) as a subspace of the action space (grey space).

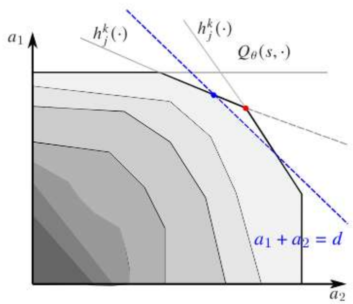


Figure 2.4: Visualization of the constraint space whose boundaries are formed by the hyperplanes $h_j^k(\cdot)$ defined by the ReLU activation functions derived from the deconstructed DNN $Q_\theta(s,\cdot)$ as a MIP formulation, for a specific state s and actions a_1 and a_2. The grey are shows the increasing value (from darker to lighter) of ∇Q_θ. The red point exemplifies the optimal solution of $\max_{a\in\mathcal{A}} Q_\theta(s,\cdot)$ if constraint $a_1 + a_2 = d$ is disregarded. If such a constraint is added to the MIP formulation, the solution represented with the blue point will be reached.

TU Delft

IEPG Intelligent Electrical Power Grids

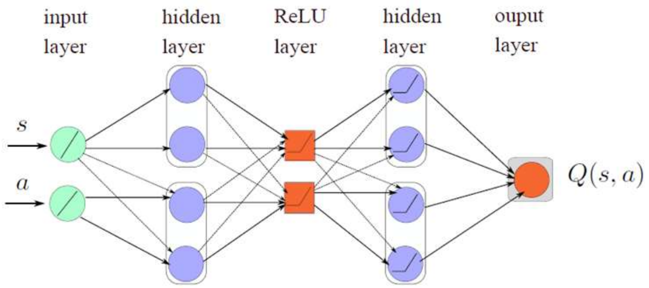
ENERGY QUANT 能源量化 Research Institution

Case Study: Energy System Optimal Scheduling

To strictly enforce the power balance constraint:

Deep neural networks + combinatorial optimization.

(MIP-DQN)



PYOMO

$$\max_{a\in\mathcal{A}, x_j^k, s_j^k, z_j^k, \forall k} \quad \{(21)\}$$

$$\left.\begin{array}{r} \sum_{i=1}^{l_{k-1}} w_{ij}^{k-1} x_i^{k-1} + b_j^{k-1} = x_j^k - s_j^k \\ x_j^k, s_j^k \geq 0 \\ z_j^k \in \{0,1\} \\ z_j^k = 1 \rightarrow x_j^k \leq 0 \\ z_j^k = 0 \rightarrow s_j^k \leq 0 \end{array}\right\} \forall k, \forall j, \quad (22)$$

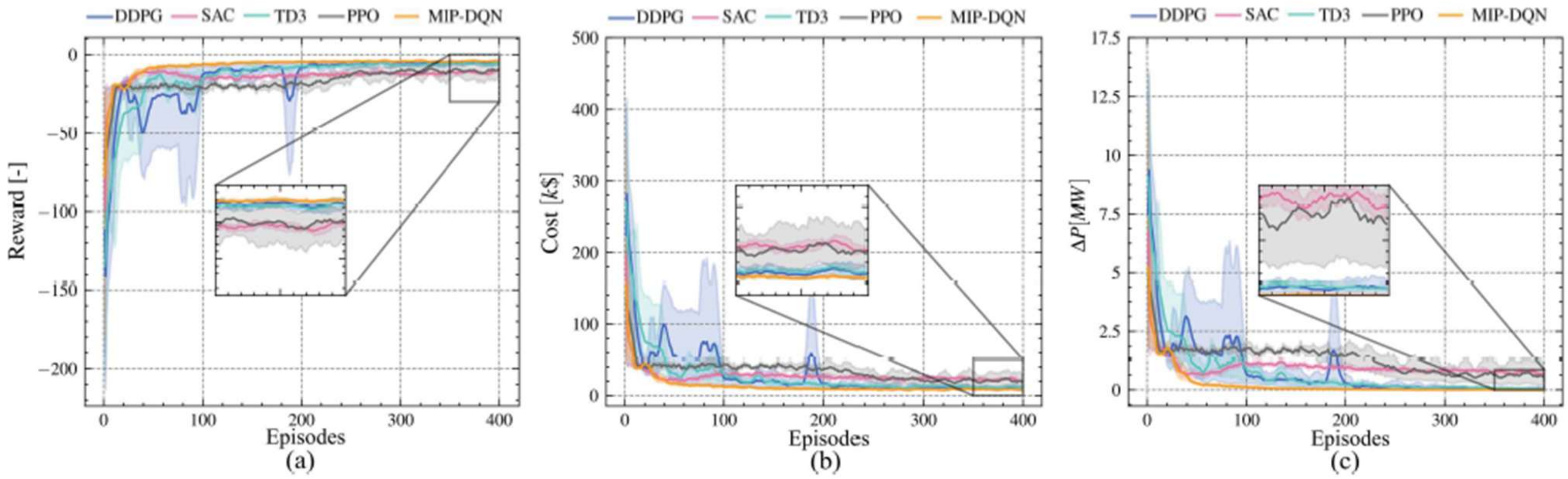
$$lb_j^0 \leq x_j^0 \leq ub_j^0, \quad j \in l_0, \quad (23)$$

$$\left.\begin{array}{r} lb_j^k \leq x_j^k \leq ub_j^k \\ \overline{lb}_j^k \leq s_j^k \leq \overline{ub}_j^k \end{array}\right\} \forall k, \forall j. \quad (24)$$

$$\sum_{i\in\mathcal{G}} P_{i,t}^G + \sum_{m\in\mathcal{V}} P_{m,t}^V + P_t^N + \sum_{j\in\mathcal{B}} P_{j,t}^B = \sum_{k\in\mathcal{L}} P_{k,t}^L, \forall t \in \mathcal{T}$$

“Deep neural networks and mixed integer linear optimization,” M. Fischetti and J. Jo, *Constraints*, vol. 23, 2018, pp. 296–309.

Case Study: Energy System Optimal Scheduling



During training, all tested algorithms seem to have similar convergence properties.

None of these algorithms are able to strictly enforce constraints, as expected. Nevertheless, the proposed MIP-DQN algorithm showed the lower error.

TU Delft IEPG

“Optimal Energy System Scheduling Using a Constraint-Aware Reinforcement Learning,” Shengren H., Pedro P. Vergara, Edgar M. Sa
IJEPES, 2022 (submitted).

ENERGY QUANT 能源量化 Research Institution

Case Study: Energy System Optimal Scheduling

Testing with unseen operational scenarios (uncertain PV and demand):

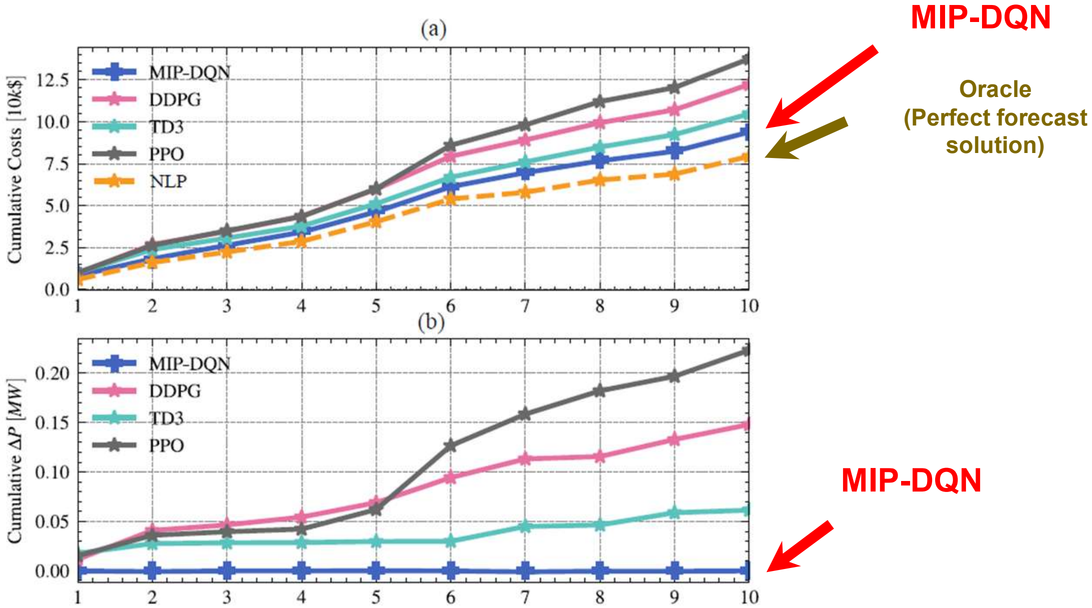


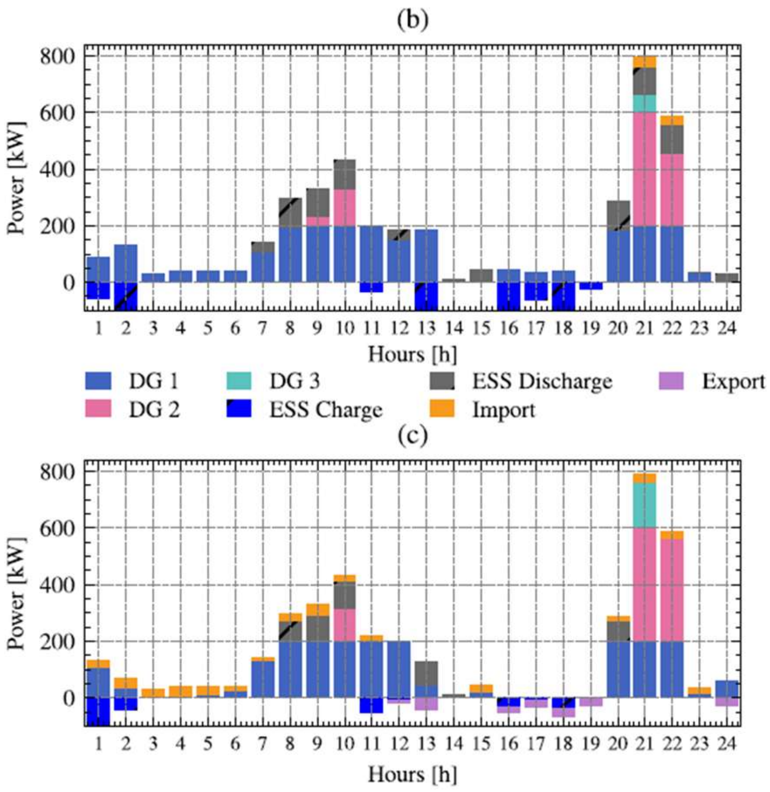
Table 4: Performance comparison of different DRL algorithms in a new test set of 30 days.

Algorithms	Error	ΔP $[MW]$	Computational time $[s]$
MIP-DQN	$13.7 \pm 0.3\%$	**0.0**	17
DDPG	47.3±1.9%	0.14±0.021	4.3
TD3	31.5±0.7%	0.06±0.011	4.9
PPO	52.4±0.3%	0.15±0.007	4.3

The MIP-DQN algorithm ***strictly*** meets the power balance constraint. Other SoA algorithms fail to do so.

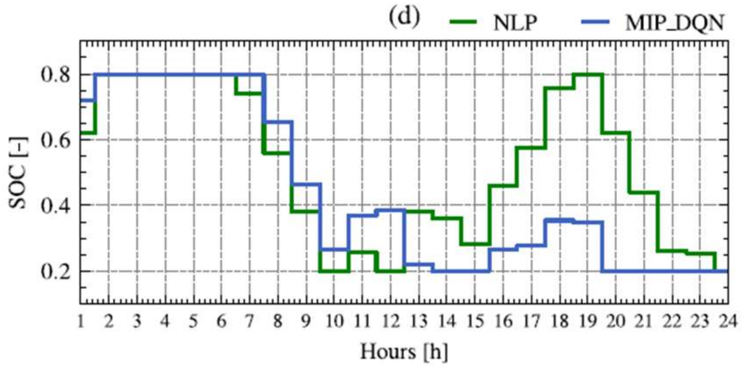
MIP-DQN algorithm achieves ***lower*** (average) errors when compared with other DRL algorithms.

Case Study: Energy System Optimal Scheduling



The MIP-DQN algorithm was able to define *similar* hourly operational schedule when compared with the optimal global solution.

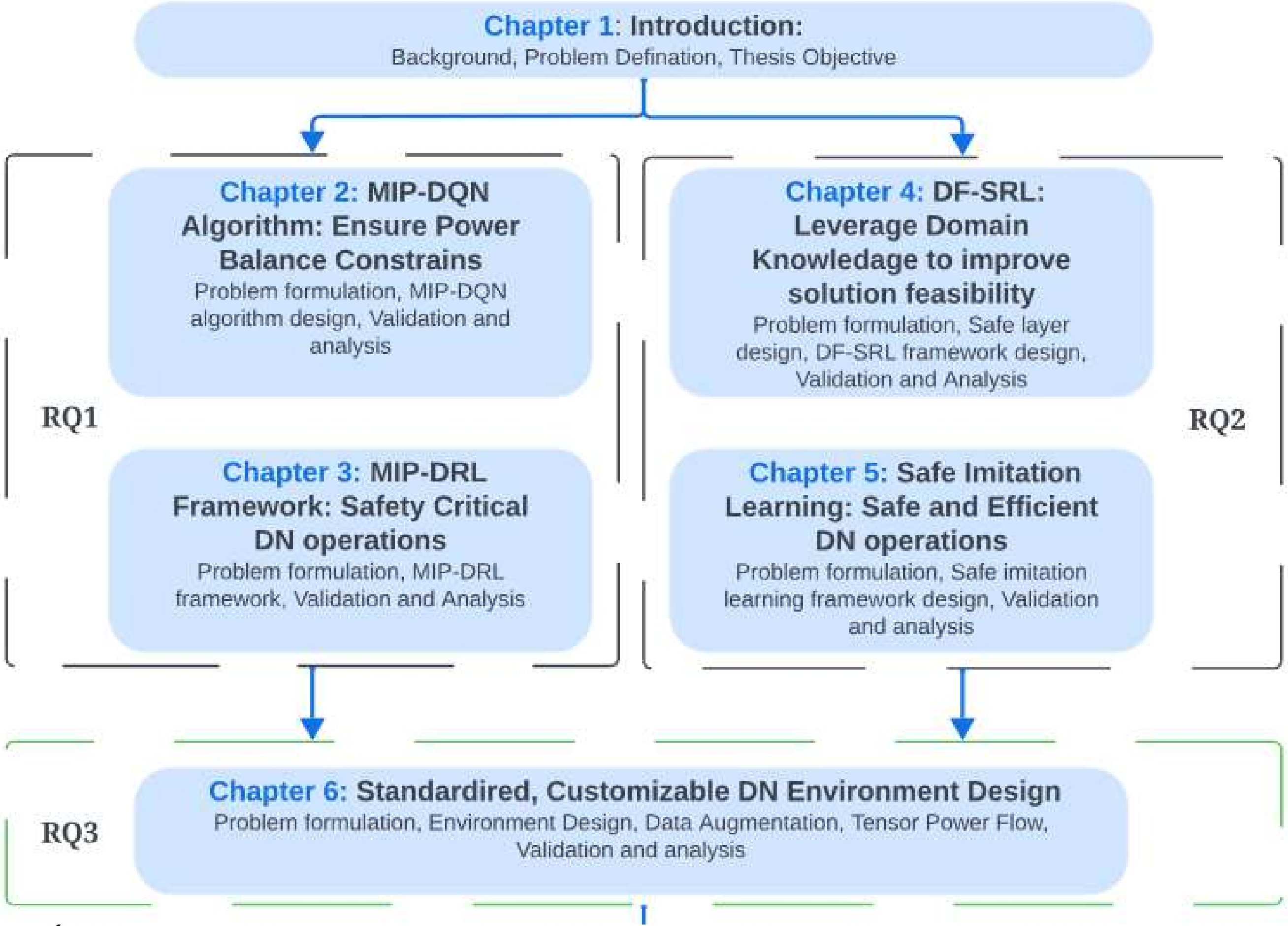
Main difference: The MIP-DQN algorithm makes decision based only on current information, while the optimal global requires estimation of future values for the stochastic variables.



Future improvement: Look better into the future. Reduce error and learn from less data (data efficiency)

TU Delft IEPG Intelligent Electrical Power Grids

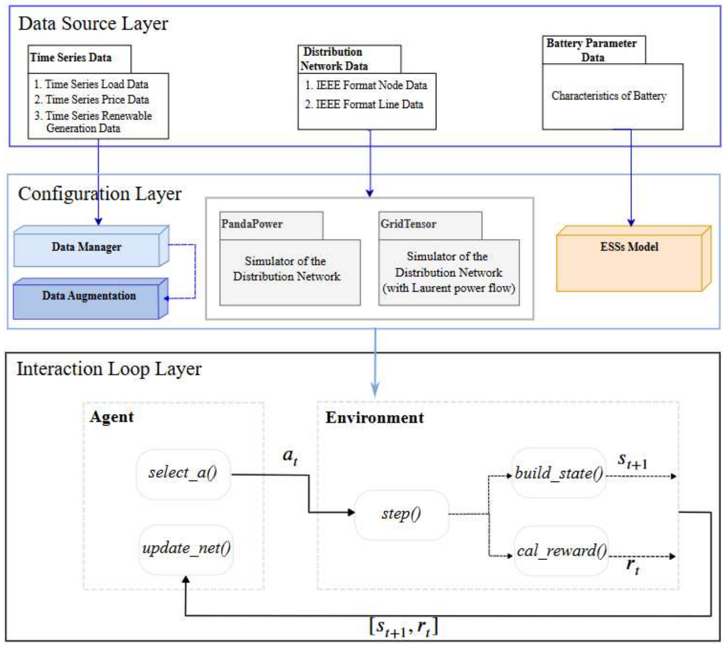
My PhD Routine



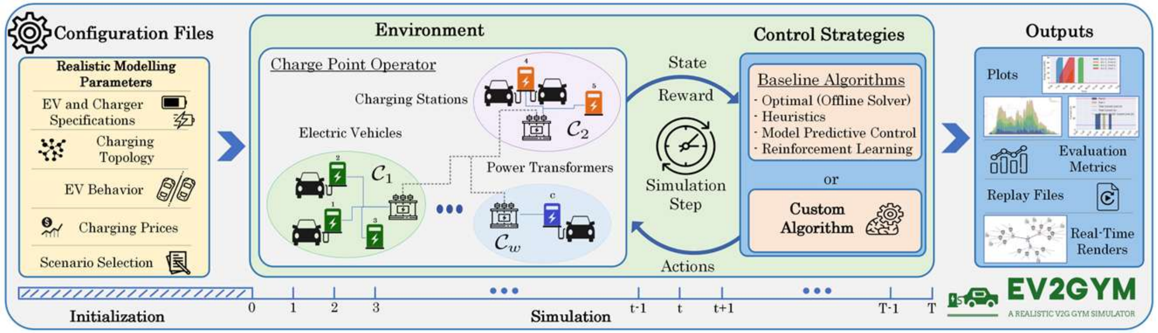
1. Stay Curious
2. Design or find out your own reward function

Some Open-source package we developed

RL-AND: An environment for ESSs dispatch in distribution network



EV2Gym: A Realistic EV-V2G-Gym Simulator for EV Smart Charging



https://github.com/ShengrenHou

https://github.com/distributionnetworksTUDelft

TU Delft

IEPG Intelligent Electrical Power Grids

Intelligent Electrical Power Grids

Distribution Networks TU Delft
distributionnetworksTUDelft

Hou Shengren
ShengrenHou

ENERGY QUANT 能源量化
Research Institution