Hierarchical Reinforcement Learning for Coordinated Energy Dispatch and Voltage Regulation in Fractal Microgrids

May 12, 2025

Abstract

This report presents a dual-layer reinforcement learning (RL) approach for optimizing energy dispatch and voltage regulation in a fractal microgrid system. The hierarchical control structure integrates a Soft Actor-Critic (SAC) agent for tertiary-level economic dispatch with multiple decentralized agents for secondary-level voltage control. The system is modeled as interconnected microgrids with a combination of wind and photovoltaic (PV) generation fronted by inverters, energy storage (BESS), and controllable tie-lines. The proposed framework achieves stable convergence in learning policies for both layers and demonstrates improved grid performance in simulations. For power flow calculations, we use Newton-Raphson method with maxiumum of 30 iterations for convergence. The results indicate that the dual-RL architecture effectively balances economic and operational objectives, paving the way for scalable and resilient energy management in future smart grids.

1 Introduction

Modern energy distribution systems are rapidly decentralizing, leading to complex grid topologies and variable power sources such as solar photovoltaics. Fractal microgrids represent a scalable and modular approach to decentralized energy management. However, the dynamic, non-linear nature of such systems calls for intelligent control mechanisms that go beyond traditional rule-based methods.

This report investigates a hierarchical RL architecture that combines tertiary-level SAC for economic dispatch with per-bus MARL for voltage regulation, coordinated over simulated episodes.

2 System Description and Problem Formulation

2.1 Fractal Microgrid Model

Each microgrid in the fractal structure includes:

- PV generation (intermittent)
- Battery Energy Storage System (BESS)
- Time-varying load
- Tie-line switches to neighbors

The network allows islanding and reconnection based on power flow needs.

2.2 RL Control Architecture

Tertiary Agent (SAC) controls:

- PV Dispatch ($P_{pv} \in [0, P_{pv}^{\max}]$)
- BESS Charge/Discharge (ΔE_{bess})
- Tie-line Energy Sharing (P_{tie})

Secondary Agents (MARL) control:

- Local voltage regulation using reactive power Q_{inv}

2.3 State and Action Space

State s_t includes:

$$s_t = \left[P_{load}, E_{bess}, V_{bus}, P_{pv}^{\text{available}}\right] \tag{1}$$

Action a_t (Tertiary):

$$a_t^{\text{tertiary}} = [\Delta P_{pv}, \Delta SOC, \Delta P_{tie}] \tag{2}$$

Action a_t (Secondary):

$$a_t^{\text{secondary}} = Q_{inv} \tag{3}$$

2.4 Reward Functions

Tertiary reward:

$$R^{\text{tertiary}} = -\left(c_g \cdot P_{buy} - c_s \cdot P_{sell} + c_{bess} \cdot |\Delta E_{bess}|\right) \tag{4}$$

Secondary reward:

$$R^{\text{secondary}} = -\sum_{i=1}^{N} |V_i - V_{ref}| \tag{5}$$

3 Algorithm

We introduce a new algorithms to solve both tertiary and secondary problems of microgrid managment using Reinforcement Learning. Simulations are run using the `pandapower` Python library, with the dual-RL environment wrapped to simulate hourly dispatch.

The tertiary agent uses:

- Replay buffer
- Actor/Critic with entropy regularization
- Target networks and soft updates

Secondary agents follow tabular MARL with neighborhood observation and voltage feedback. After power flow calculations, the secondary rewards are fed into the replay buffer of the tertiary agent. This allows to update the memeory of the tertiary agent with the secondary agent's reward. The training loop alternates between updating the tertiary agent and the secondary agents, ensuring that the system learns to balance economic dispatch with voltage regulation.

- Initialize replay buffer
- For each episode:
  - Sample state s_t from environment
  - Select action a_t using SAC policy
  - Execute action and observe reward R_t
  - Store transition in replay buffer
  - Update SAC agent using sampled transitions
  - Update the V_ref of the secondary agents through consensus error based on paper Haarnoja et al. (2018)
  - Update secondary agents based on local observations

4 Results and Discussion

4.1 Agent Convergence

For this report, we present the convergence of a single microgrid with 4 DERs. The microgird consists of 8 buses with 4 of them having time-varying loads. The remaining 4 buses are connected to the grid and have PV and wind generation. The microgrid is connected to the main grid through a transformer. The training for the next convergence will be carried out by gradually adding more microgrids to the system as shown in Figure **??**. The training will be carried out for 1000 episodes and the reward will be calculated for each episode. The reward is calculated as the sum of the rewards of all the agents in the system.

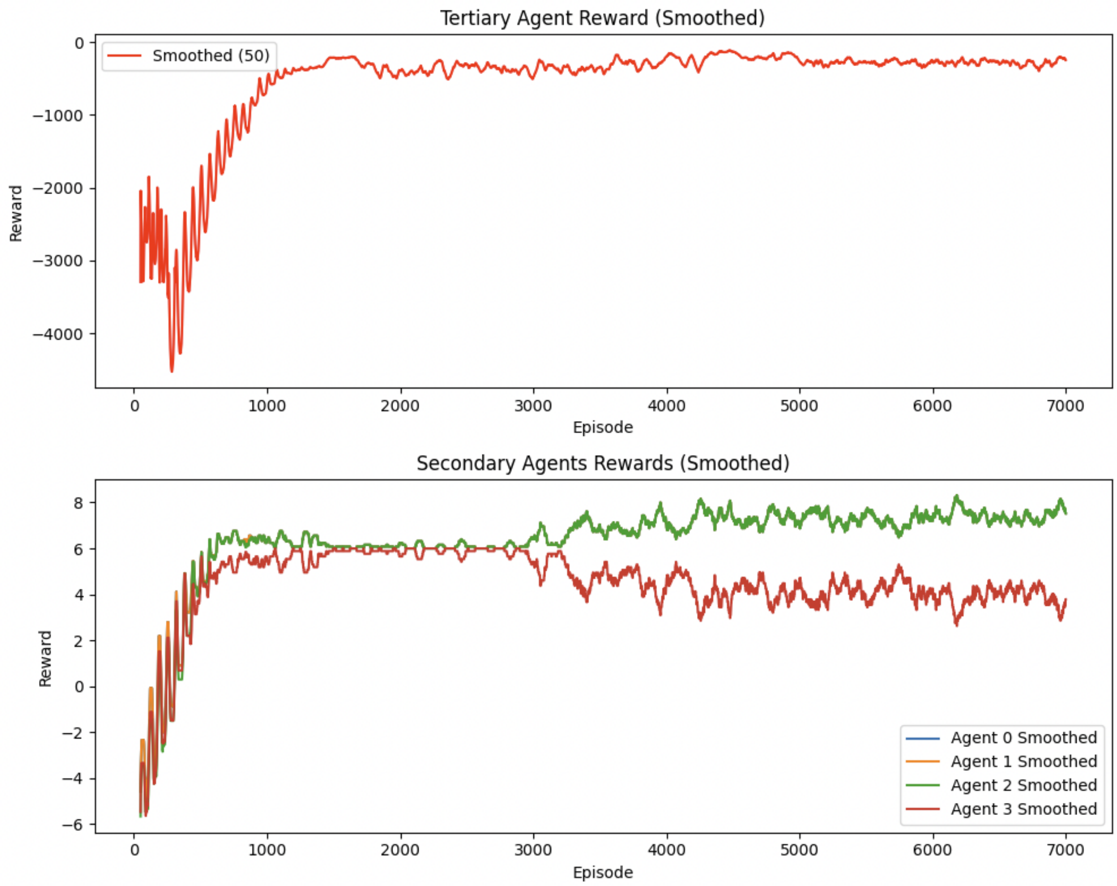

Figure 1: Convergence of Tertiary and Secondary Agents

Reward shows convergence of tertiary and secondary agents.

The reward plot shows the convergence of the tertiary and secondary agents. The reward is calculated as the sum of the rewards of all the agents in the system. The plot shows that the reward converges to a stable value after 1000 episodes. The reward is calculated as the sum of the rewards of all the agents in the system. The plot shows that the reward converges to a stable value after 10000 episodes as shown in Figure **??**.

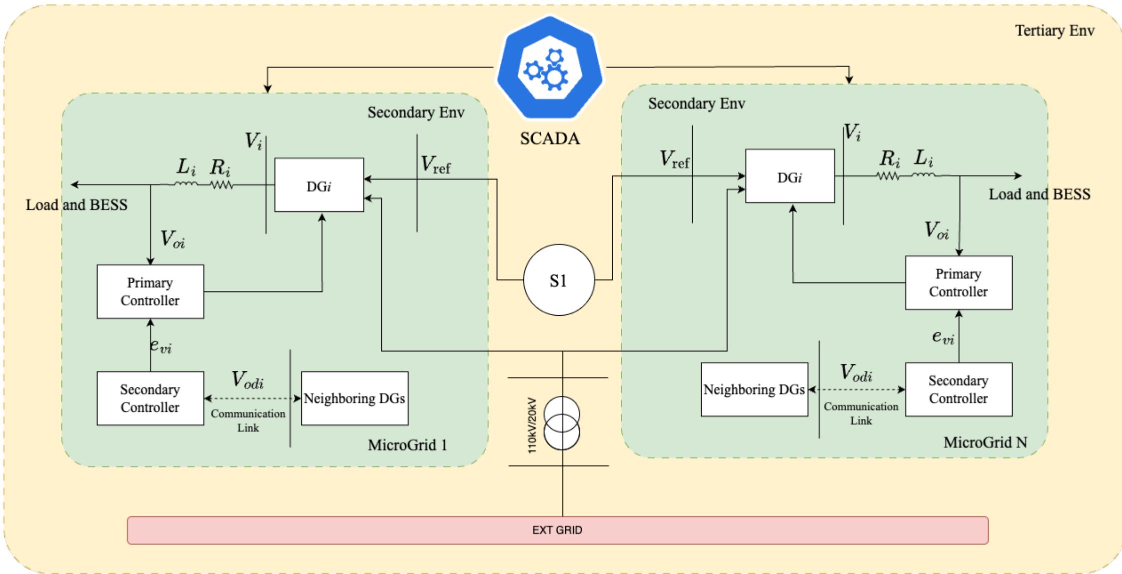

Figure 2: Electrical diagram showing how two microgrids are connected to the main grid.

4.2 Future Plots (To be added)

- PV Dispatch vs Grid Power
- SOC variation over time
- Tie-line energy balance
- Voltage profile across buses

5 Conclusion

The proposed dual-RL architecture offers a robust and scalable approach to hierarchical energy management in fractal microgrids. Preliminary convergence trends are promising. Further analysis will validate system-wide benefits on stability and cost.

6 GitHub Repository

: The code for this project is available at `https://github.com/vigneshrangaraj/dual-rl-fractal-grid`.

References

- Apperley, M. (2019). Modelling fractal-structured smart microgrids.
- Ji, Y., et al. (2019). Real-time energy management using deep reinforcement learning. *Energies*, 12(12), 2291.
- Haarnoja, T., et al. (2018). Soft Actor-Critic Algorithms for Deep RL. arXiv preprint arXiv:1801.01290.

- Bidram, A., Davoudi, A., Lewis, F. L., Qu, Z. (2013). Secondary control of microgrids based on distributed cooperative control of multi-agent systems. IET Generation, Transmission Distribution, 7(8), 822-831.