Energy Saving for Cell-Free Massive MIMO Networks: A Multi-Agent Deep Reinforcement Learning Approach

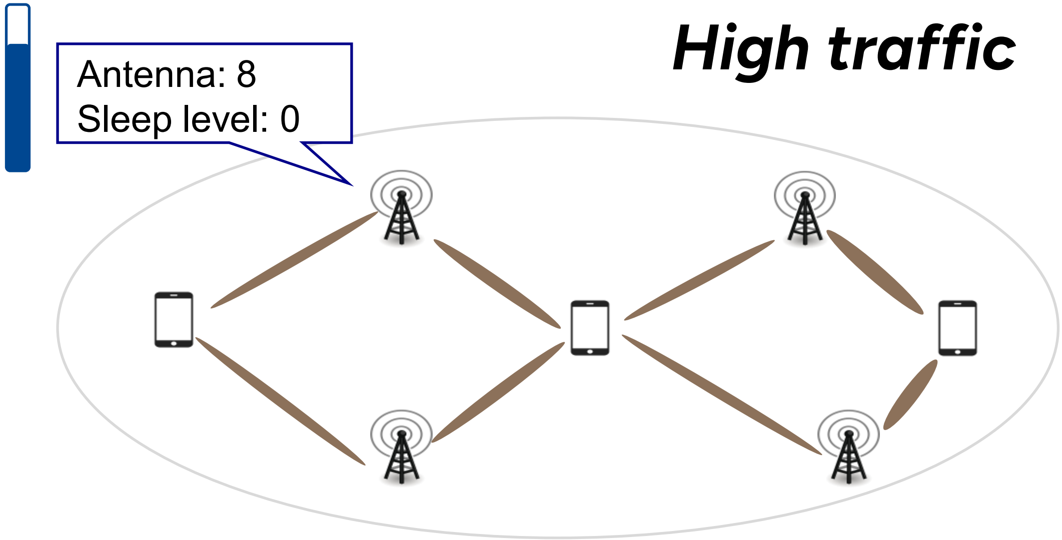
Qichen Wang [1], Keyu Li [1], Ozan Alp Topal [1], Özlem Tugfe Demir [2], Mustafa Ozger [3], Cicek Cavdar [1]

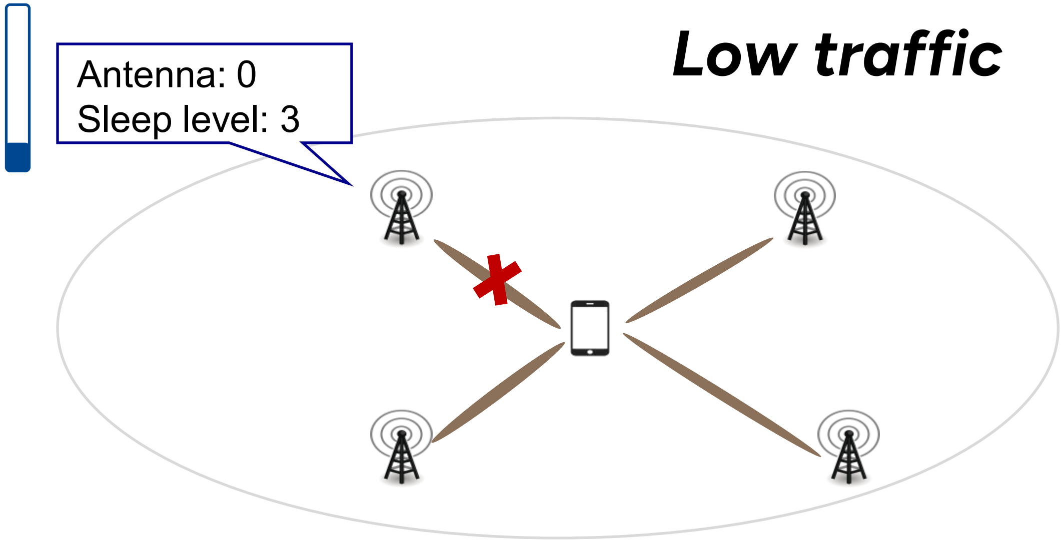
[1] Department of Computer Science, KTH Royal Institute of Technology, Stockholm, Sweden

[2] Department of Electrical and Electronics Engineering, Bilkent University, Ankara, Türkiye

[3] Department of Electronic Systems, Aalborg University, Denmark

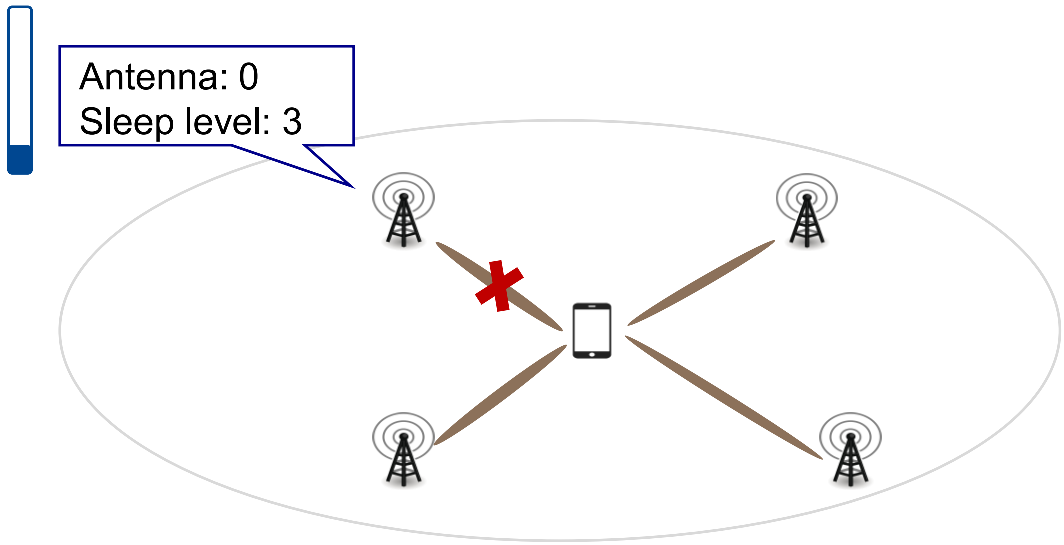
Introduction / Motivation





- Cell-free massive MIMO is a promising architecture for future networks, where many distributed APs jointly serve users to provide more uniform QoS.
- In low-traffic periods, many APs and antennas may remain unnecessarily active, leading to wasted energy.
- Energy can be saved by deactivating antennas or switching APs to sleep modes.
- This motivates adaptive AP control: each AP should adjust its antenna configuration and sleep mode according to traffic variations.

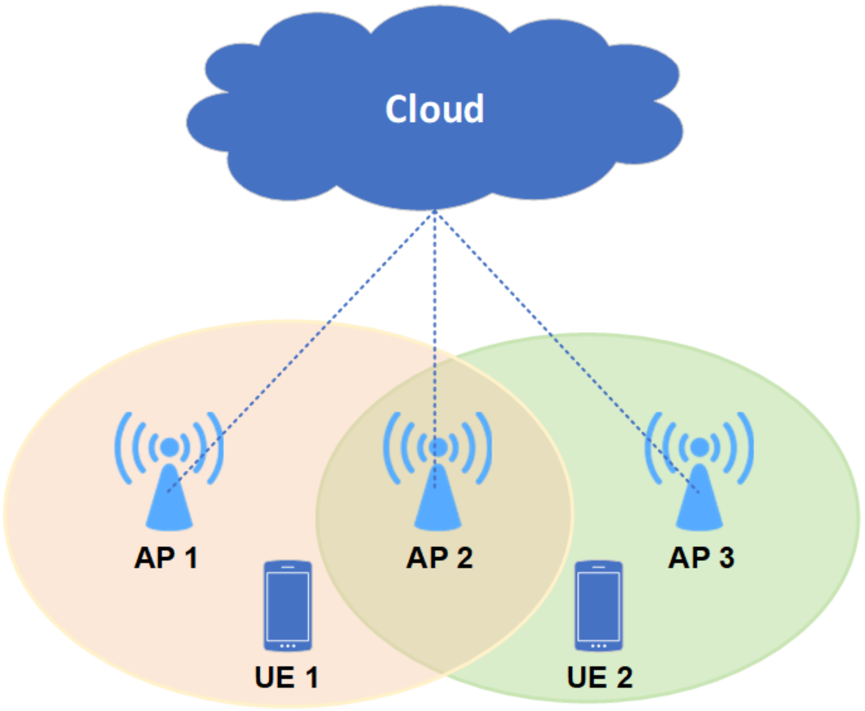
Contribution



Goal: *Jointly optimize AP antenna configuration and sleep mode selection under dynamic traffic to reduce power consumption while ensuring QoS.*

- We develop a digital twin of a dynamic CF mMIMO system by using real mobile traffic data from a Swedish operator to emulate time-varying UE arrivals.
- We propose a MAPPO (multi-agent proximal policy optimization)-based algorithm.

System Model



Network model:

- We consider the downlink of a CF mMIMO network.
- The network consists of L distributed APs, each equipped with up to M_{max} antennas.

Association:

- Each UE is jointly served by a subset of distributed APs selected based on the strongest large-scale channel gains.

Precoding:

- Local protective partial zero-forcing (PPZF) precoding.

Control:

- Each AP can adapt its active antennas and sleep mode as traffic changes.

Power Model

$$P_l = m_l P_{\mathrm{st}} + \Delta_{\mathrm{tr}} \sum_{k \in \mathcal{K}_l} p_{l,k} + \left(P_0^{\mathrm{proc}} + \Delta_{\mathrm{AP}}^{\mathrm{proc}} \frac{C_{\mathrm{AP},l}}{C_{\mathrm{AP}}^{\max}} \right)$$

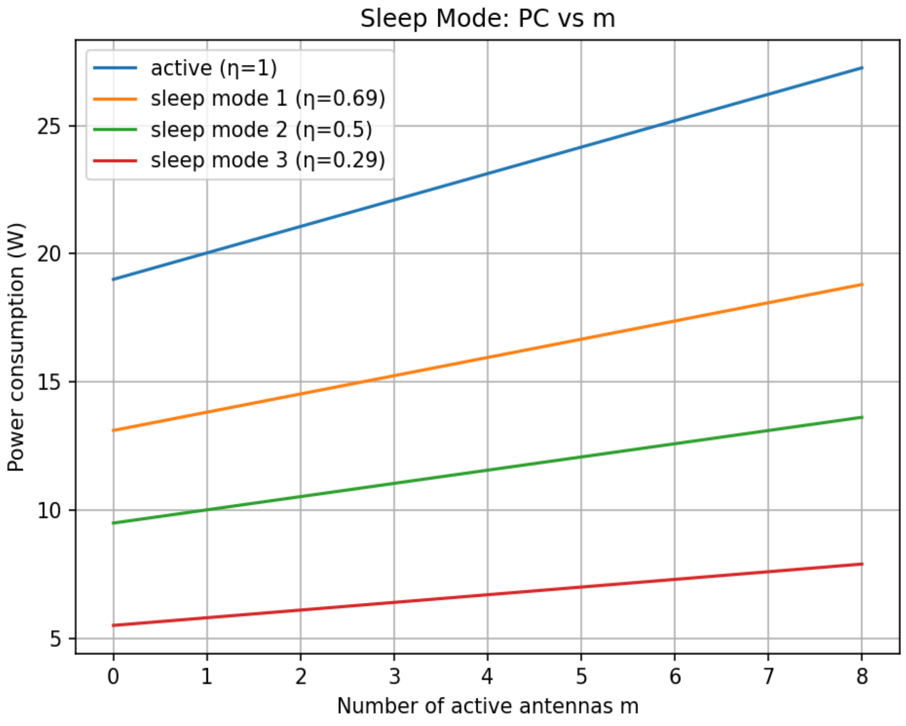
Active mode:

- Power consumption increases with the number of active antennas and served UEs.

Sleep mode:

- Deeper sleep modes consume less power, but require a longer wake-up time.

Sleep mode s	0	1	2	3
Wake-up latency Δ_s	$0\ \mu s$	$37\ \mu s$	$500\ \mu s$	$5000\ \mu s$
PC discount factor η_s	1	0.675	0.55	0.23

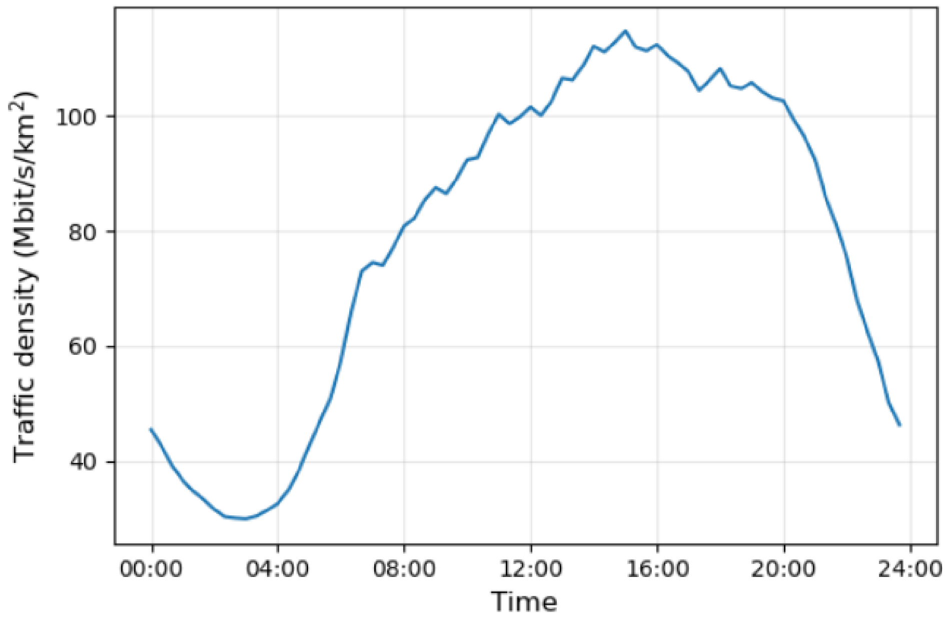


Traffic Model

- UE arrivals follow a space-homogeneous time-varying Poisson process.

$$\lambda = \frac{\kappa A \Delta t}{x^{max}}$$

- The arrival intensity is derived from operator DPI traffic data.
- Sessions are aggregated into 20-minute intervals to derive the temporal-spatial traffic density κ.



- Each UE has a traffic demand x^{max} and delay constraint D^{max} based on 3GPP specifications.
- A UE leaves the system once its demand is fully served or its delay budget runs out.

Optimization Problem

$$
\begin{aligned}
\min_{s_l,\, m_l,\, \forall l} \quad & P_{\text{net}} \\
\text{s.t.} \quad & \frac{1}{K}\sum_{k=1}^{K} \frac{x_k^{\text{rem}}}{x^{\max}} \leq \delta_{\text{drop}}, && \forall k \\
& m_l \in \{0, 1, \ldots, M_{\max}\}, && \forall l \\
& \tau_l^{\text{str}} \in \{0, 1, \ldots, m_l - 1\}, && \forall l \\
& s_l \in \{0, 1, 2, 3\}, && \forall l
\end{aligned}
$$

Objective:

- Minimize the total power consumption.

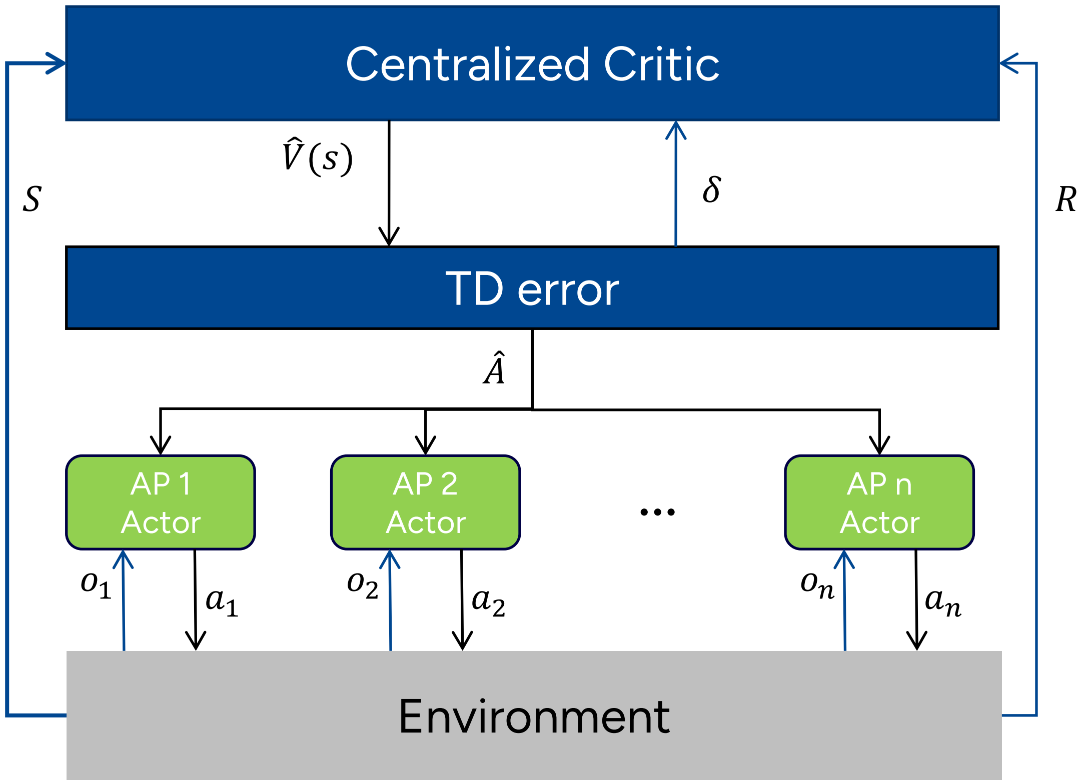
Decision variables per AP:

- Sleep mode s_l and antenna activation m_l.

Constraints:

- Average drop ratio.
- Number of active antennas.
- Non-zero effective SINR for PPZF precoding.
- Available sleep modes.

Multi Agent Proximal Policy Optimization (MAPPO)



Algorithm structure:

- Each AP is modeled as an agent.

Centralized training, decentralized execution:

- Actors use local observations to select antenna and sleep-mode actions.
- The centralized critic evaluates joint actions using the global state during training.
- After training, only the actors are deployed at the APs.

RL Algorithm Design

Local obervation O:

- AP state: power, active antennas, and sleep mode.
- Aggregated UE statistics: demand and rate.
- Neighboring AP states.

Global state S:

- Total network power.
- UE arrival rate.
- Aggregated UE statistics.

Action A:

$A = A_m \times A_s$

- $A_m = \{-1, 0, 1\}$: antenna switching
- $A_s = \{0, 1, 2, 3\}$: sleep level

Reward R:

A shared reward balances QoS and PC.

$$R = w_{qos} \frac{1}{K} \sum_{k=1}^{K} \xi_k - w_{pc} P_{net}$$

- QoS of UE: $\xi_k = \begin{cases} \rho_k - 1, & \rho_k < 1 \\ \phi\left(1 - \frac{1}{\rho_k}\right), & \rho_k \geq 1 \end{cases}$
- $\rho_k = \frac{actual\ throughput\ of\ UE\ k}{required\ rate\ of\ UE\ k}$

Simulation Setup

Key settings:

- 1 km x 1 km area; 25 APs on a 5 x 5 grid with 8 antennas
- 20 MHz bandwidth at 5 GHz
- 1 ms simulation steps; actions every 20 ms
- MAPPO trained for 200 episodes

Benchmarks:

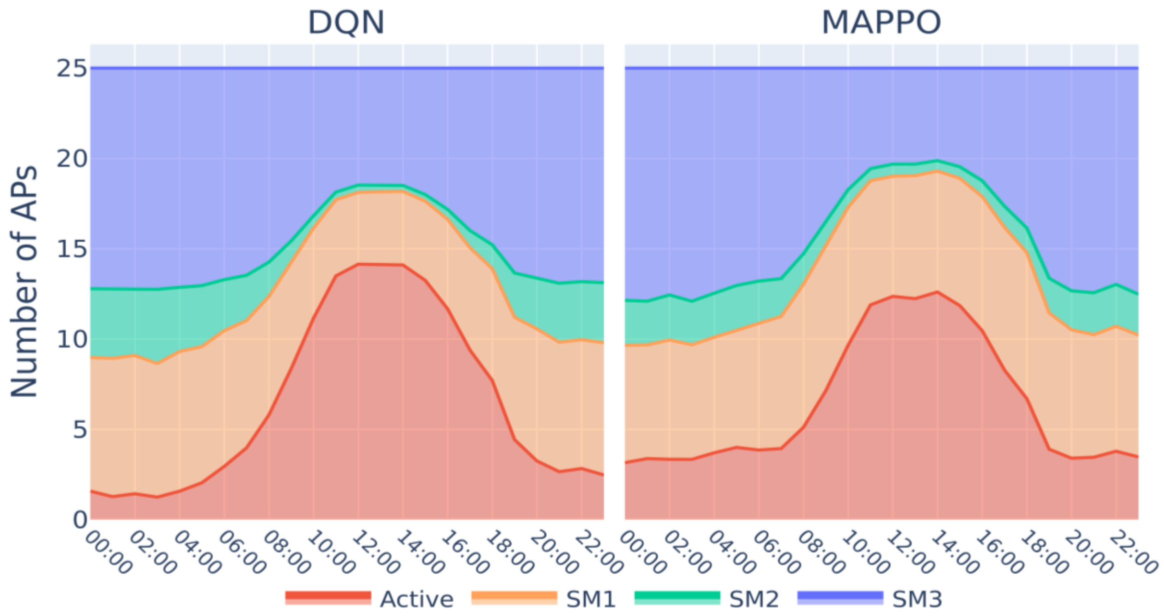
- **Always-on:** all APs and antennas remain active.
- **DAC-SM1:** SM1 when idle plus threshold-based antenna control.
- **DQN:** Deep Q-learning benchmark.

Metrics:

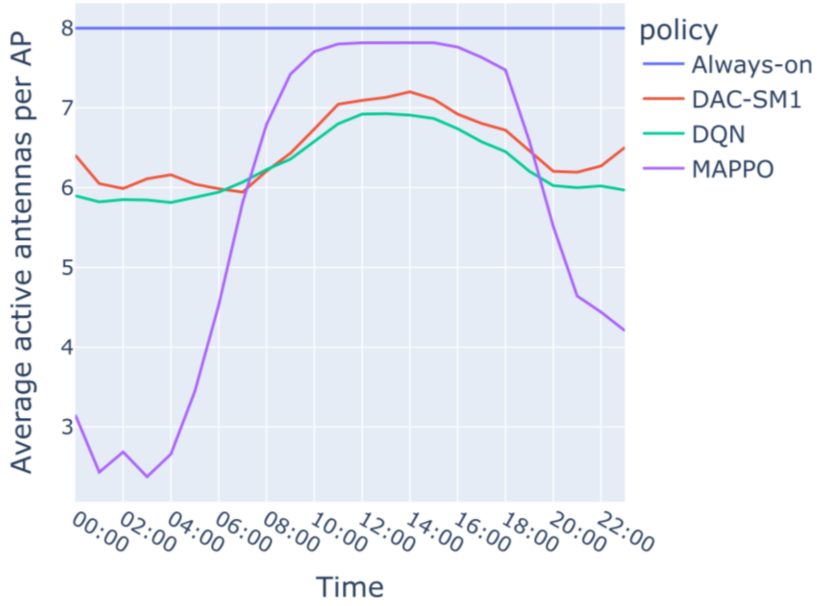
- Network PC
- Average UE drop ratio.

Numerical Results

Time-varying number of APs in different sleep modes:



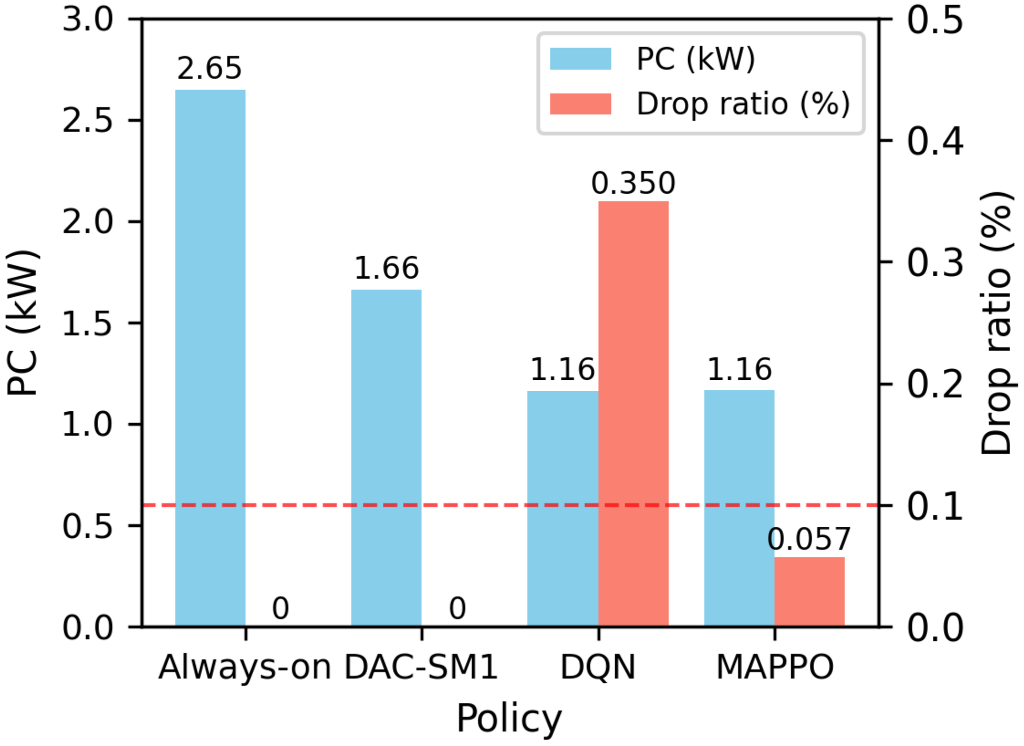
Average number of active antennas per AP:



- **MAPPO** dynamically adapts sleep modes and antenna activation to traffic variations.
- **MAPPO** achieves a much wider adjustment range (2 to 8 antennas).
- **DQN** and **DAC-SM1** exhibit limited flexibility in antenna control.

Numerical Results

Overall performance:



- **MAPPO** achieves up to 56% PC reduction compared to **Always-on** and 30% compared to **DAC-SM1**.
- **MAPPO** maintains a very low drop ratio (well below the dash line constraint).
- **DQN** achieves similar power but suffers from a much higher drop ratio.

Conclusion and Future Work

Conclusion:

- Digital twin of a dynamic CF mMIMO system with real mobile traffic data.
- MAPPO-based algorithm for antenna activation and sleep modes.
- Power consumption falls by 56.23% vs. **Always-on** and 30.12% vs. **DAC-SM1**, with only a slight increase in drop ratio.

Future work:

- Currently, all UEs are treated equally. By differentiating UEs according to their QoS requirements, the network can avoid unnecessarily activating antennas or APs, thereby achieving further energy savings.

Thanks for your attention

qichenw@kth.se