
DTU

Reinforcement Learning Control of Raman Amplifiers
05.10.2025

REINFORCE Algorithm

- The REINFORCE algorithm is based on the weight update which is informed by the reward function.
- The weights in our case are the probabilities in the Bernoulli distribution.
- We compare the reward to the baseline and multiply that with the learning rate and eligibility
- Eligibility is a measure of how representative the current sample is of the probability

delta probs = lr * (reward – baseline) * eligibility

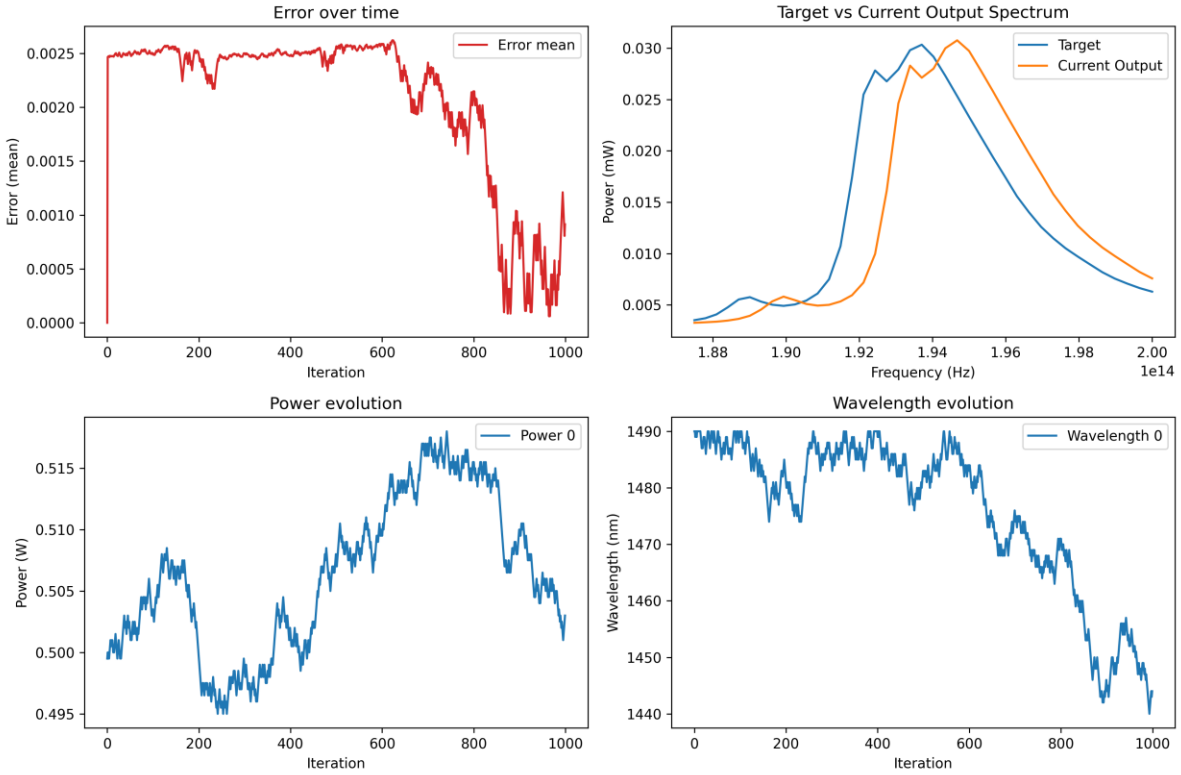
eligibility = sample – probs

baseline = gamma * baseline + (1 - gamma) * reward

sample ~Bernoulli(probs)

REINFORCE Algorithm - Results

- Some experiments resulted in good final results but no significant convergence was achieved.
- There was no noticeable pattern in the hyperparameters which resulted in good final control.
- This could be due to the fact that the system is too complex for the simple Bernoulli controller to manage to control it, or that the reward function is not adequate, or even that the hyperparameters were not selected correctly.



REINFORCE Algorithm - Results