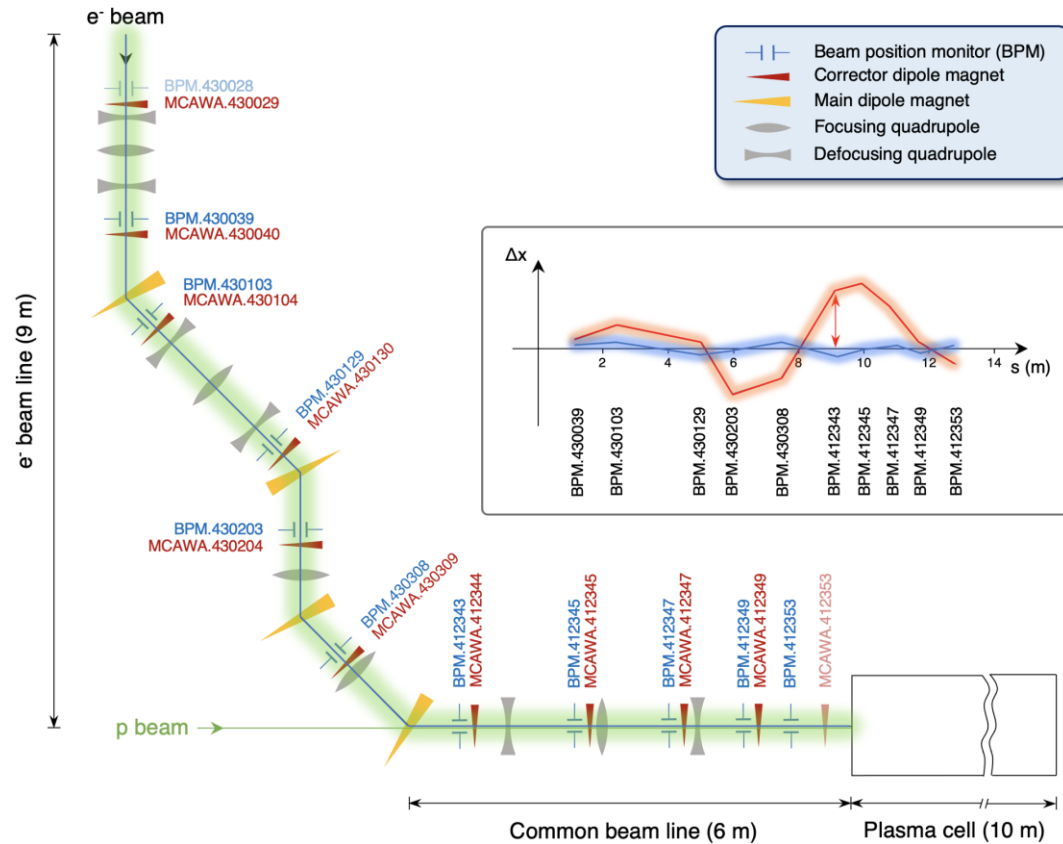


Algorithms

Example: actor-critic for AWAKE



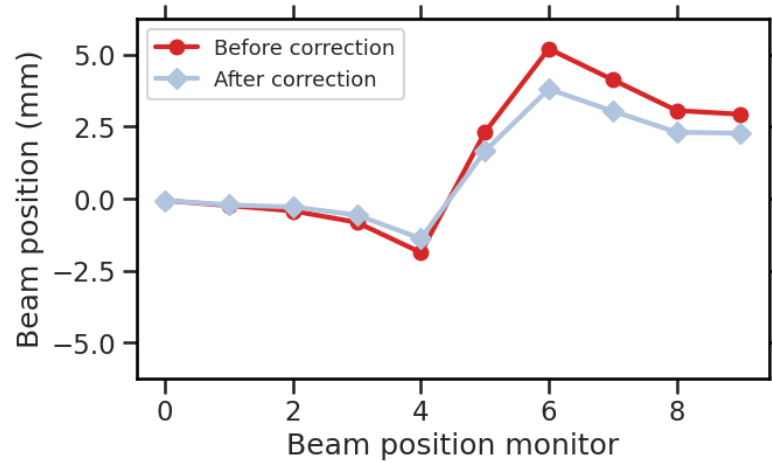
RL setting

- **Goal**
 - Given **beam position** readings, adjust **dipole correctors** to keep beam close to the centre of vacuum pipe
 - Ideally trained agent should do **single-shot** correction from any initial state
- **State:** 10-d array of beam positions measured along the line (*BPM. ...*)
- **Action:** 10-d array of dipole corrector strengths along the line (*MCAWA. ...*)
- **Reward:** negative root-mean-square (rms) of beam offsets wrt. vacuum pipe centre
- **Episodic task:** every episode starts with random initial dipole corrector settings

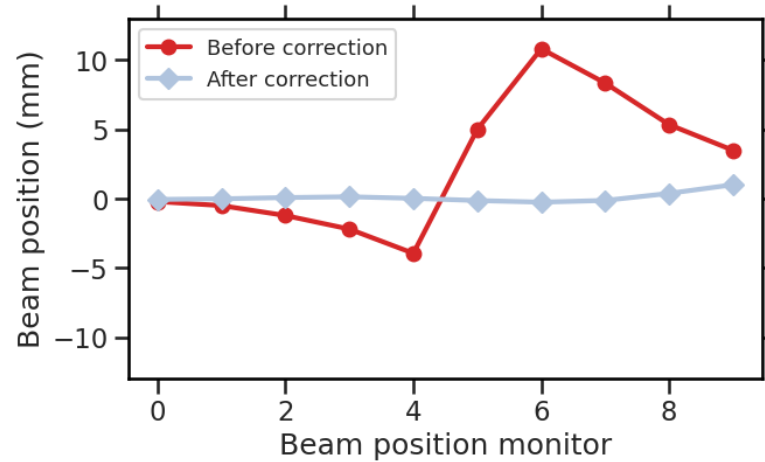
Algorithms

Example: actor-critic for AWAKE

Before training



After training



Agent training with TD3

