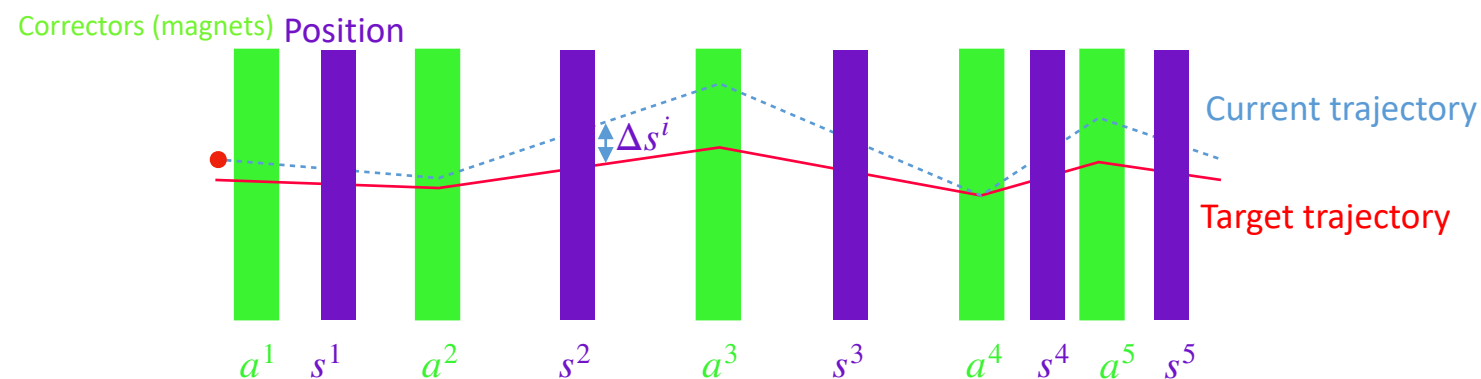


CERN AWAKE steering problem

Markov decision process: $(\mathcal{S}, \mathcal{A}, \mathcal{R}, \mathcal{P}, \rho_0, \gamma)$



- 10 continuous states \mathcal{S} and actions $\mathcal{A} \in [-1, 1]$ (**actions are bounded/constraints**) (10 DoF problem - observation is state)

- Rewards \mathcal{R} negative of RMS of states $r_t \propto -\sqrt{\sum_k \Delta(s_t^k)^2}$

- The dynamics of the system is characterised by: $s_{t+1} = \mathbf{R}a_t + s_t$
- Initial criteria: Initial distribution ρ_0
- Episodic training
- Termination criteria:
 - Maximal number of interactions (truncation)
 - RMS below measurement uncertainty (successful termination)
 - States $s_i >$ beam pipe (termination or clipping)
- Transitions \mathcal{P} are deterministic or stochastic, $\gamma = 1$
- If we speak about different tasks i (MPDs) we mean different matrices \mathbf{R}_i

Details of the problem I