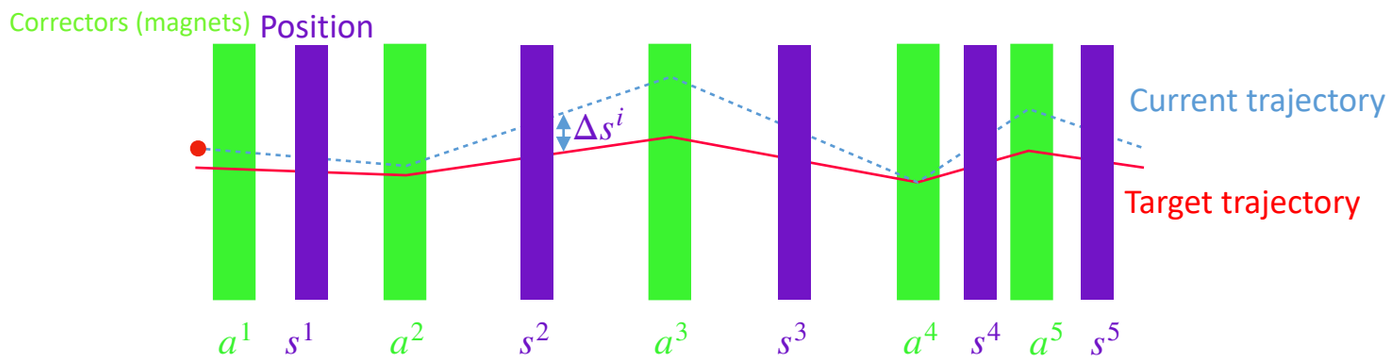
CERN AWAKE steering problem

Markov decision process: (S, A, R, P, ρ_0, γ)



- 10 continuous states S and actions $A \in [-1,1]$ **(actions are bounded/constraints)** (10 DoF problem - observation is state)
- Rewards 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 P are deterministic os stochastic, γ =1
- If we speak about different tasks i (MPDs) we mean different matrices $\mathbf{R}_i$