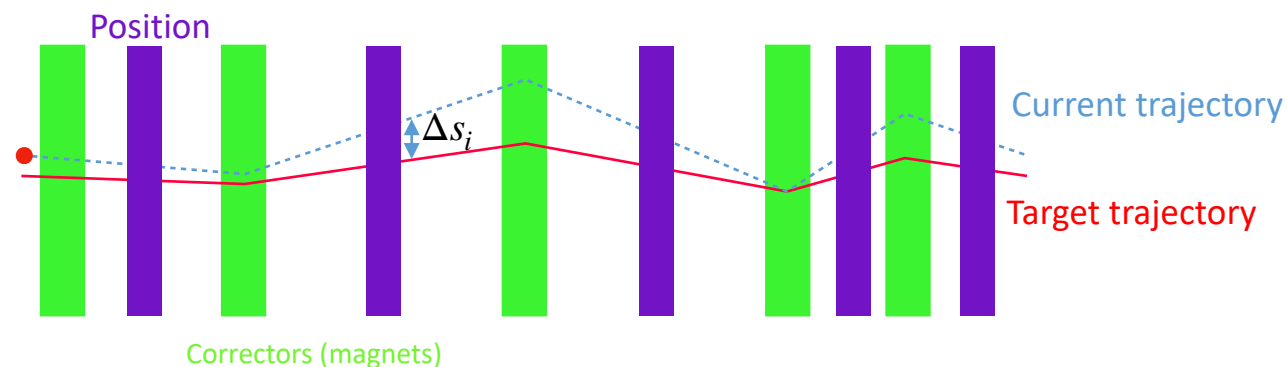


# CERN AWAKE steering problem

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



- 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  $r_i \propto -\sqrt{\sum \Delta s_i^2}$
- Actions are done in  $s_{t+1} = \mathbf{R}a_t + s_t$
- Episodic training
- Initial criteria: Initial distribution  $\rho_0$  is away from low RMS - to make problem a bit challenging
- 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 or stochastic,  $\gamma=1$
- If we speak about different tasks  $i$  (MPDs) we mean different matrices  $\mathbf{R}_i$

# Motivation