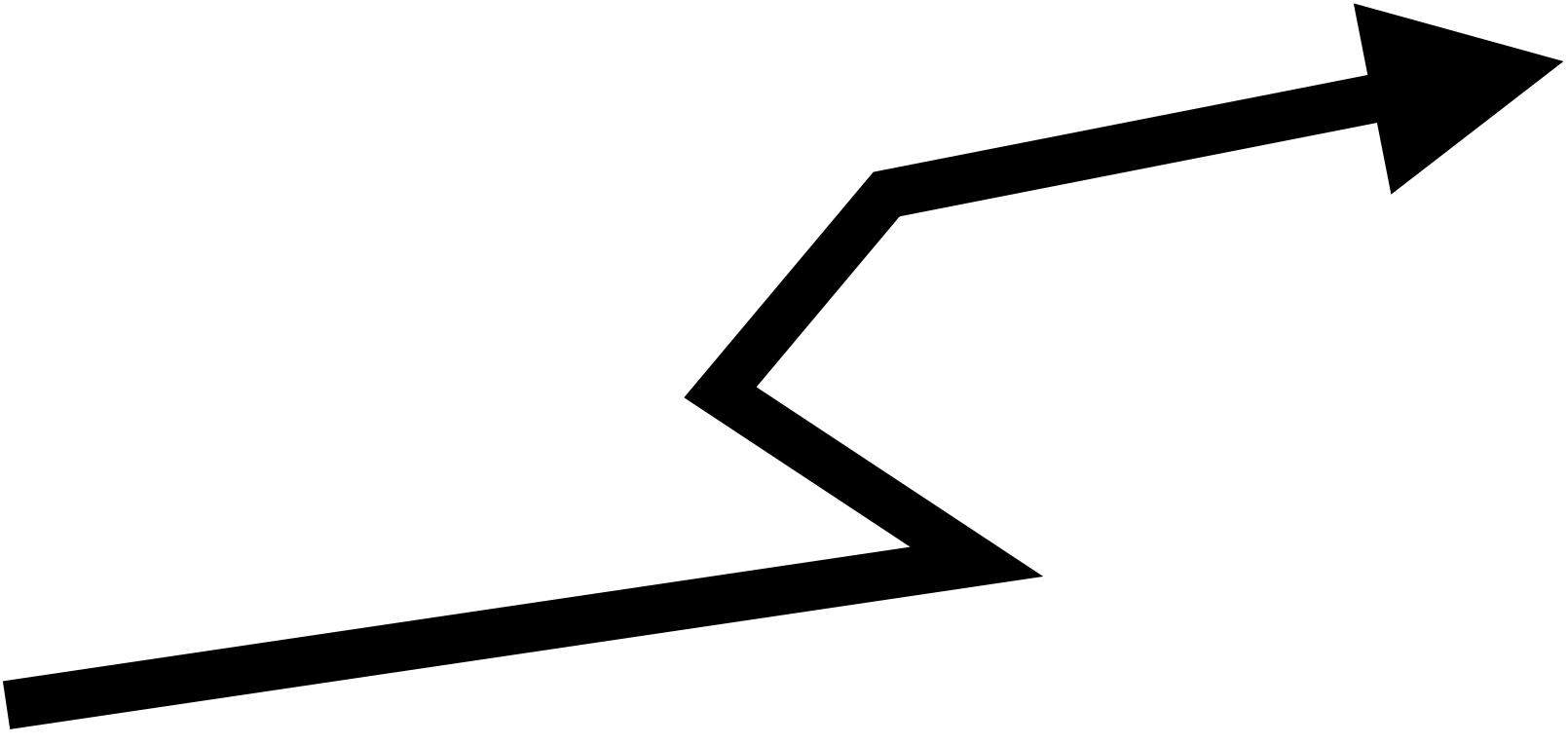
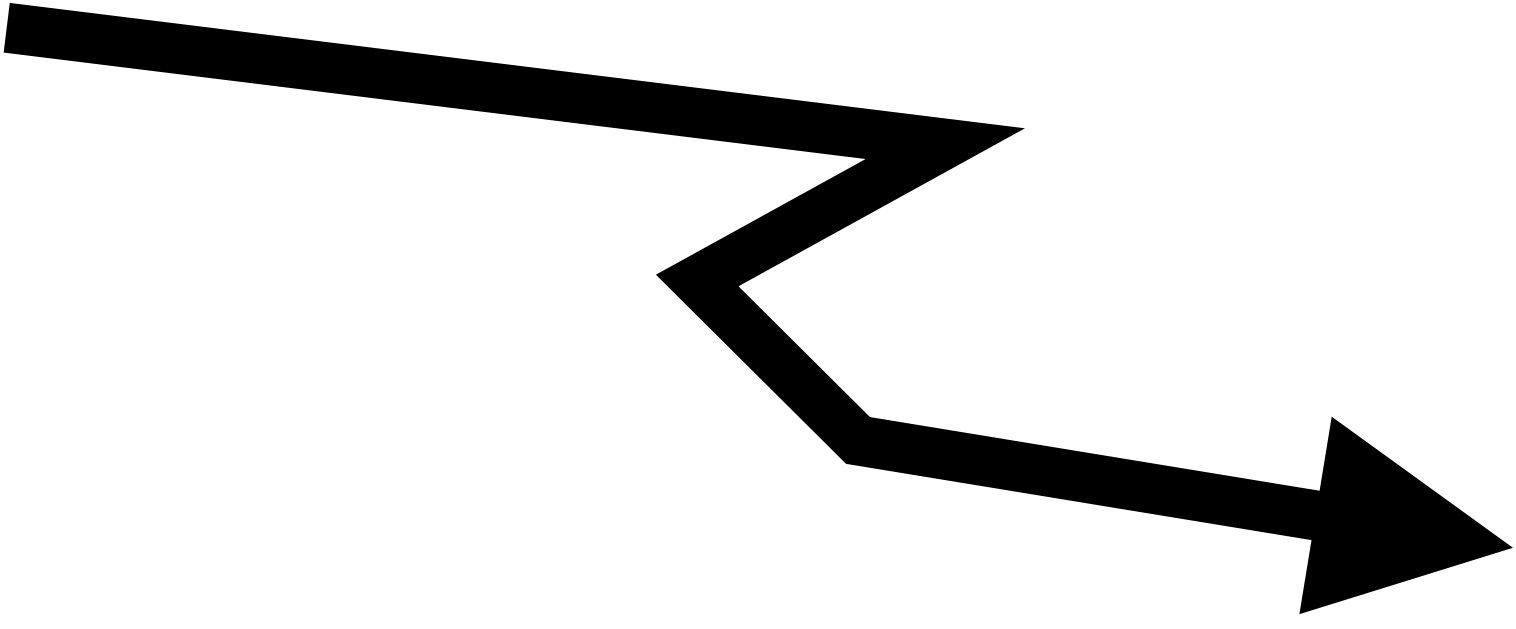
Shooting Methods

Cross-Entropy Method

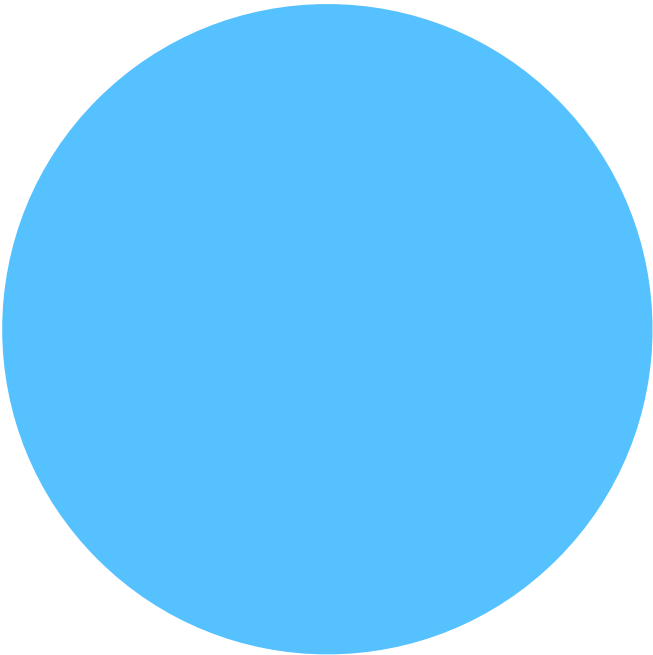
25





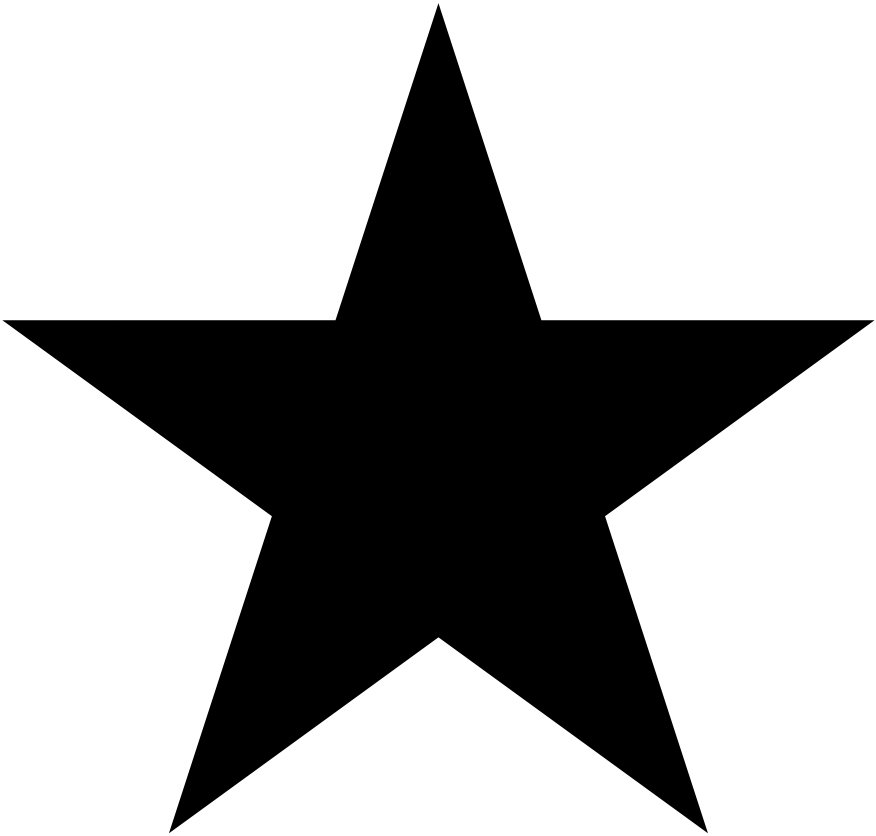
$a^*_{0:H}$

Iteration 3



More sample efficient

Faster convergence



Initialise action sequence sampling distribution $\{a_t \sim \mathcal{N}(\mu_t, \sigma_t^2)\}_{t=0}^{H}$

For each iteration

Sample N action sequences $\{a_{0:H}^i\}_{i=1}^{N}$ from sampling distribution

Evaluate objective $J(a_{0:H}^i) = \sum_{t=0}^{H} \gamma^t r(s_t, a_t^i)$ for each sample

Select top K performing samples, i.e. highest value $J(a_{0:H}^i)$

Update parameters $\{\mu_t, \sigma_t^2\}_{t=0}^{H}$ of action dist. using top K samples

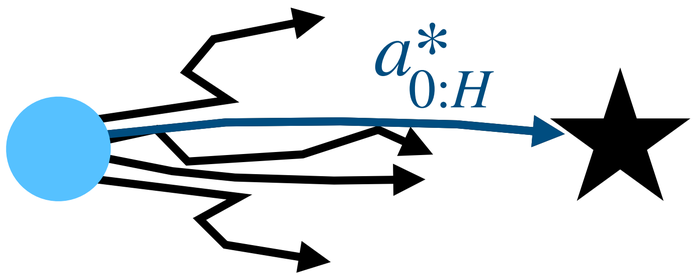
Shooting Methods

Cross-Entropy Method

More sample efficient

Faster convergence

Iteration 3



Initialise action sequence sampling distribution $\{a_t \sim \mathcal{N}(\mu_t, \sigma_t^2)\}_{t=0}^{H}$

For each iteration

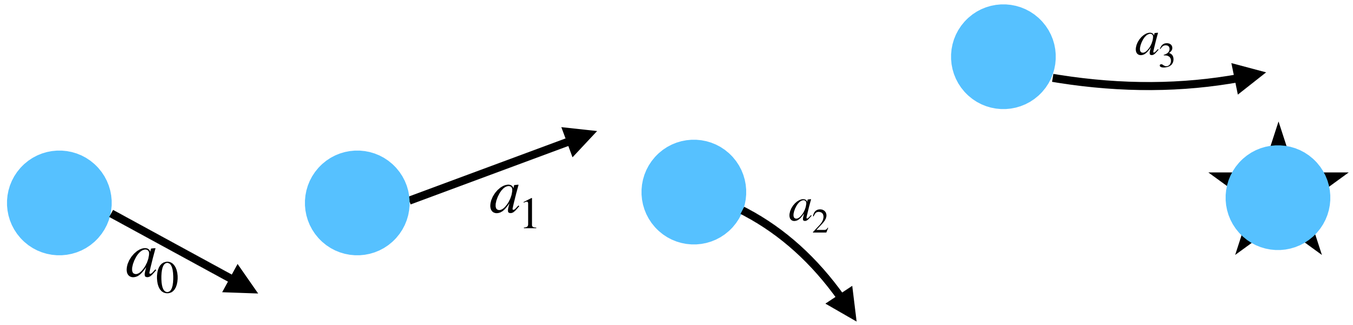
Sample N action sequences $\{a^i_{0:H}\}_{i=1}^{N}$ from sampling distribution

Evaluate objective $J(a^i_{0:H}) = \sum_{t=0}^{H} \gamma^t r(s_t, a^i_t)$ for each sample

Select top K performing samples, i.e. highest value $J(a^i_{0:H})$

Update parameters $\{\mu_t, \sigma_t^2\}_{t=0}^{H}$ of action dist. using top K samples

Collocation methods

Illustration

$$J(a_{0:H}, s_{0:H}) = \sum_{t=0}^{H} \gamma^t r(s_t, a_t) \quad \text{s.t. } \|s_{t+1} - f(s_t, a_t)\| = 0$$