



**FCAI**

**fcai.fi**

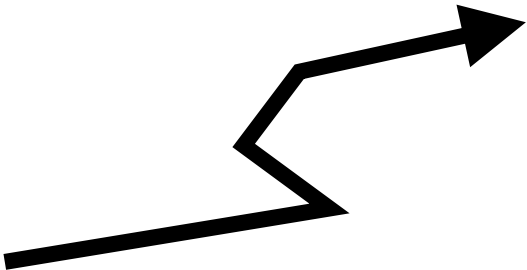
# Shooting Methods

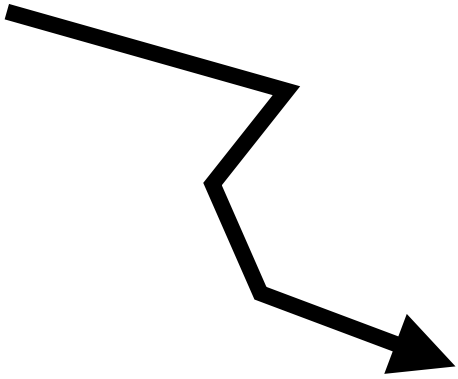
# Cross-Entropy Method

2

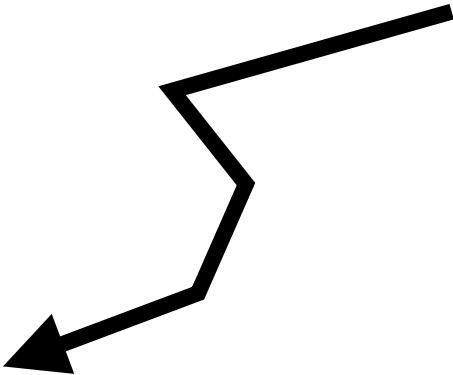
3

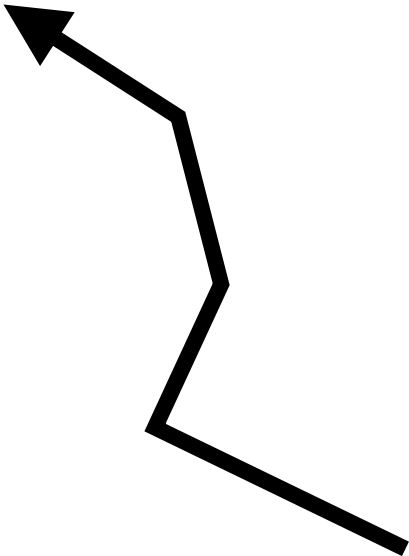


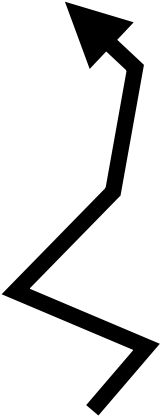








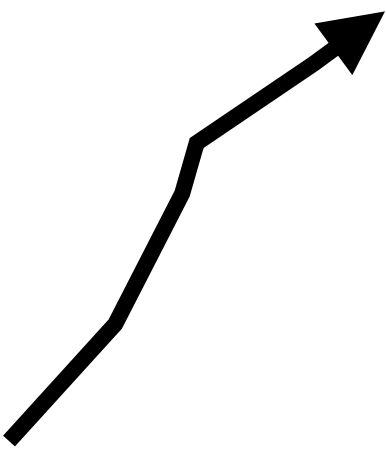






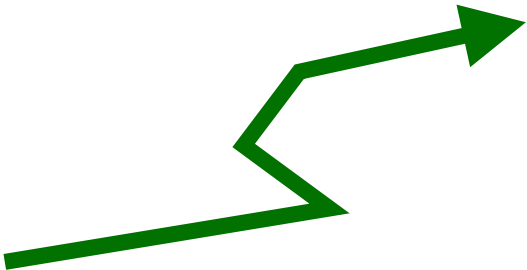


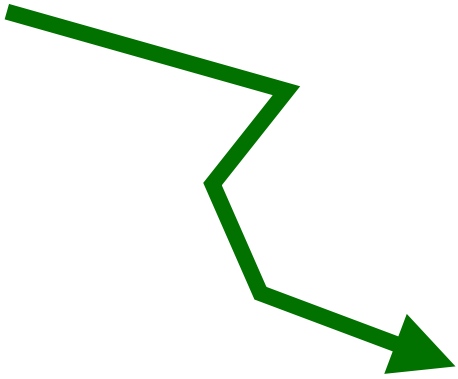




top-*K*











Iteration 1

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

For 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

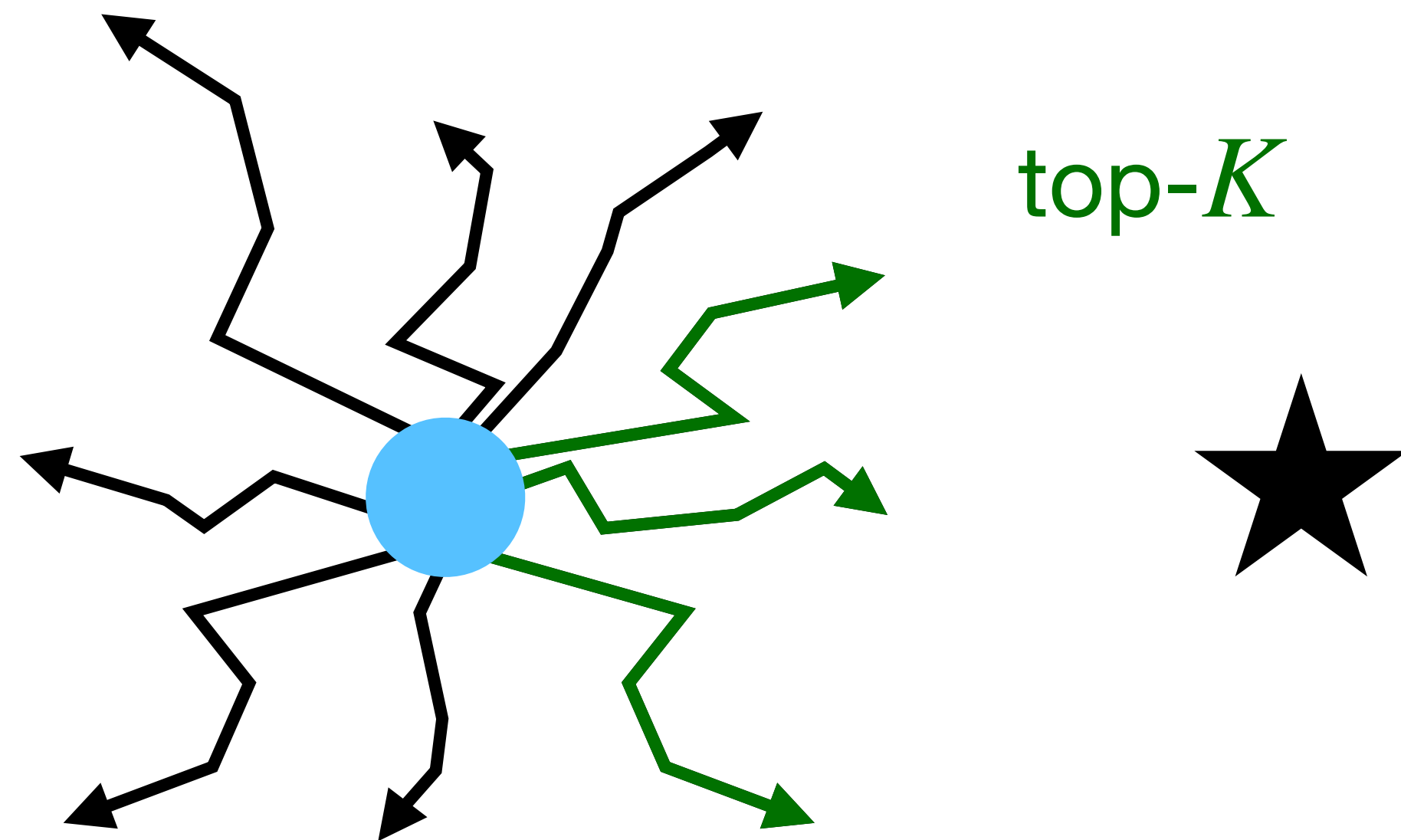
Select top  $K$  performing samples, i.e. highest value  $J(a_{0:H}^i)$

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

# Shooting Methods

## Cross-Entropy Method

Iteration 1



**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

Iteration 2

**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