

Avg reward per trajectory  $w=[0,1]$

