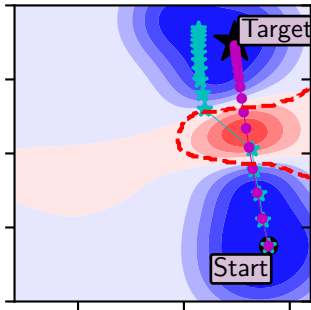
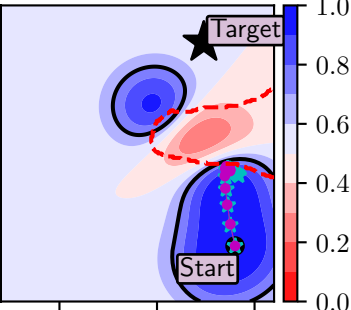


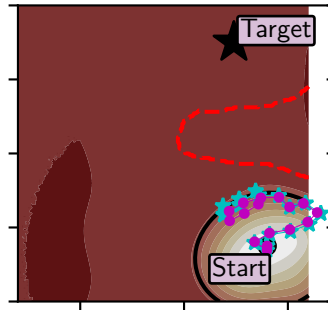
Greedy no constraint



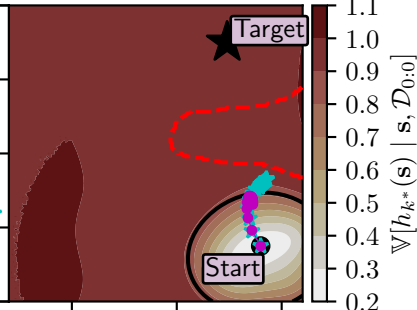
Greedy


 $\Pr(\alpha = k^* | \mathbf{s}, \mathcal{D}_{0:10})$

Non-myopic exploration



Myopic exploration



—★— Environment \bar{s}_{π_i}
 —●— Dynamics $\bar{s}_{\pi_i}^{f_{k^*}}$
 - - - Mode boundary
 — δ -mode constraint