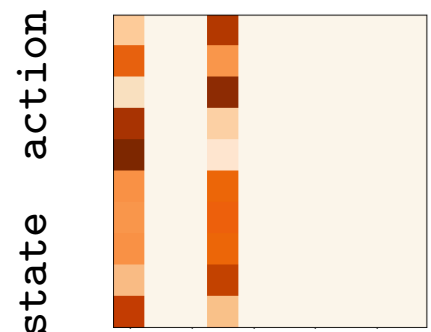


Baseline

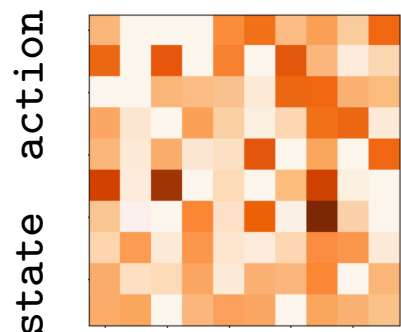
Attention map



state action

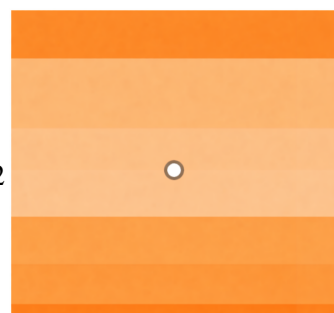
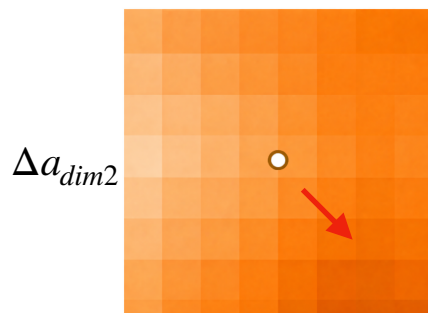
TQL

Attention map



state action

TQL prevents attention collapse

 $Q(s, a)$  $Q(s, a)$ 

Resulting in a smoother
Q-value landscape

