

KL Divergence Control

Purpose of KL Divergence

KL divergence measures how much the optimized policy deviates from the original base model, preventing harmful divergence from foundational knowledge.

Visual Representation



$$\text{KL}(\pi_\theta \parallel \pi_{\text{ref}})$$



‡ KL measures distribution difference

