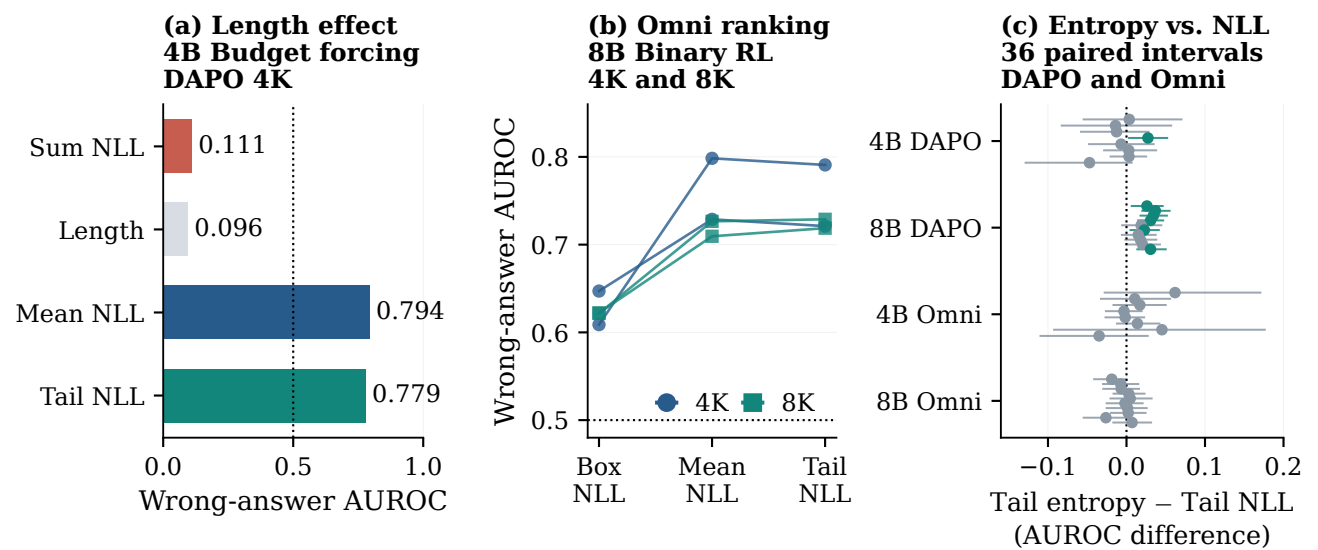

(a) Length effect
4B Budget forcing
DAPO 4K
Sum NLL
0.111
Length
0.096
Mean NLL
0.794
Tail NLL
0.779
0.0
0.5
1.0
Wrong-answer AUROC
(b) Omni ranking
8B Binary RL
4K and 8K
Wrong-answer AUROC
0.8
0.7
0.6
0.5
4K
8K
Box NLL
Mean NLL
Tail NLL
(c) Entropy vs. NLL
36 paired intervals
DAPO and Omni
4B DAPO
8B DAPO
4B Omni
8B Omni
−0.1
0.0
0.1
0.2
Tail entropy − Tail NLL
(AUROC difference)