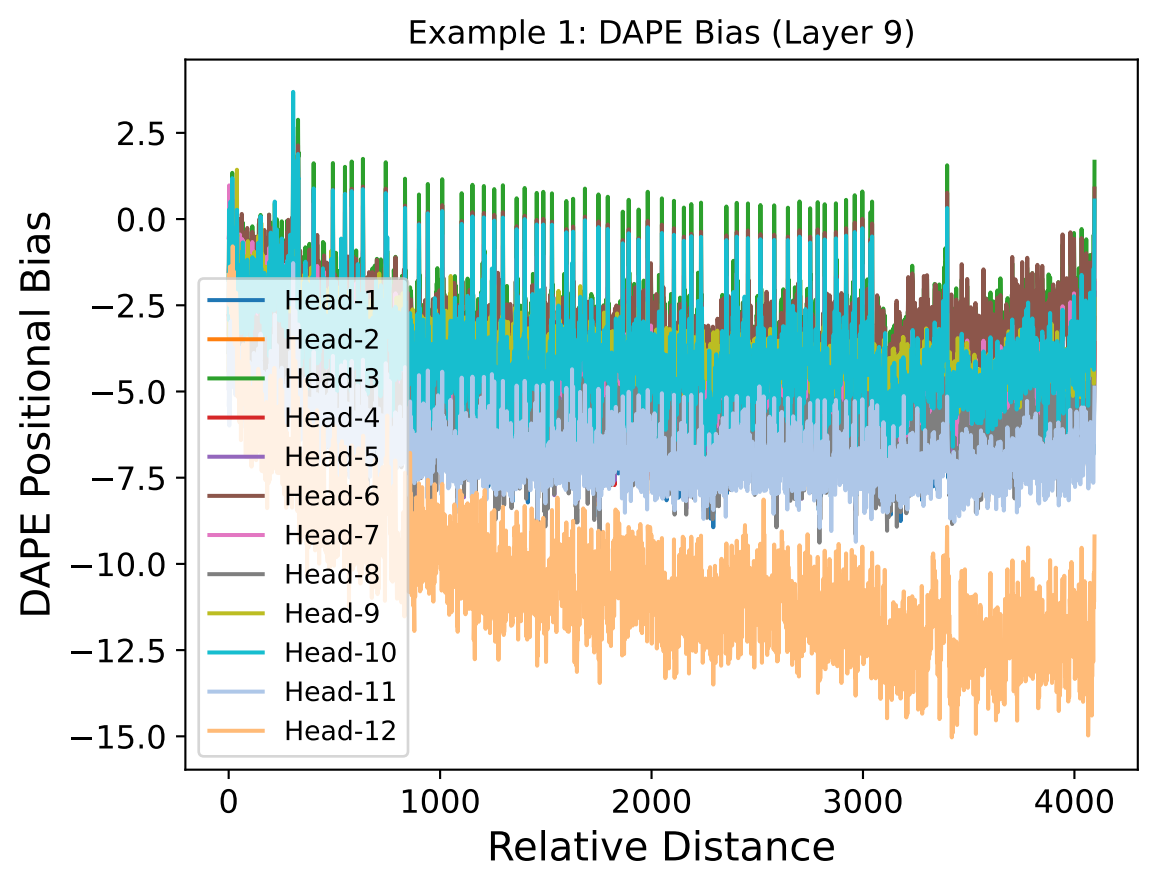
Example 1: DAPE Bias (Layer 9)
DAPE Positional Bias
Relative Distance
2.5
0.0
−2.5
−5.0
−7.5
−10.0
−12.5
−15.0
0
1000
2000
3000
4000
Head-1
Head-2
Head-3
Head-4
Head-5
Head-6
Head-7
Head-8
Head-9
Head-10
Head-11
Head-12