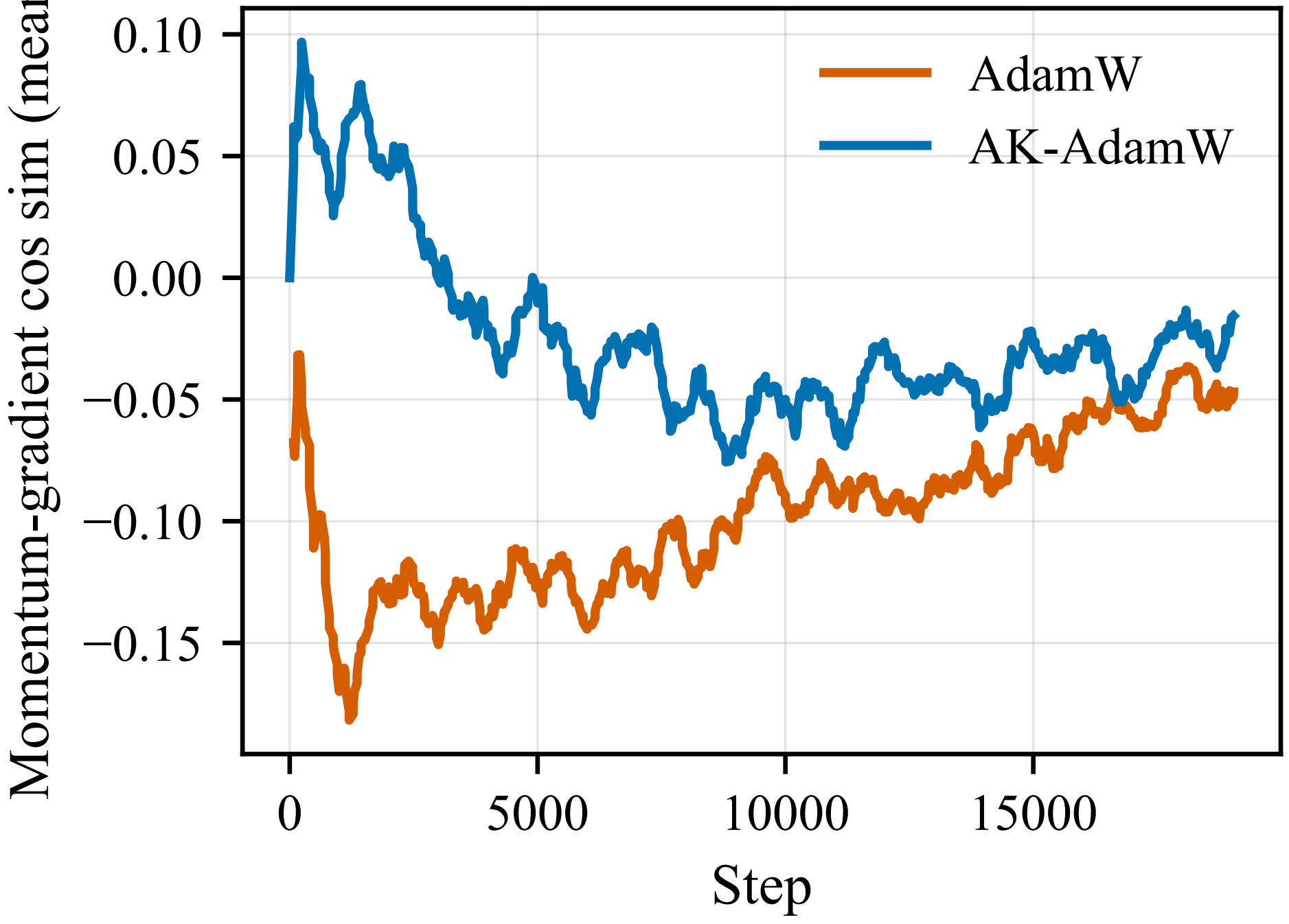

Momentum-gradient cos sim (mean
0.10
0.05
0.00
−0.05
−0.10
−0.15
0
5000
10000
15000
Step
AdamW
AK-AdamW