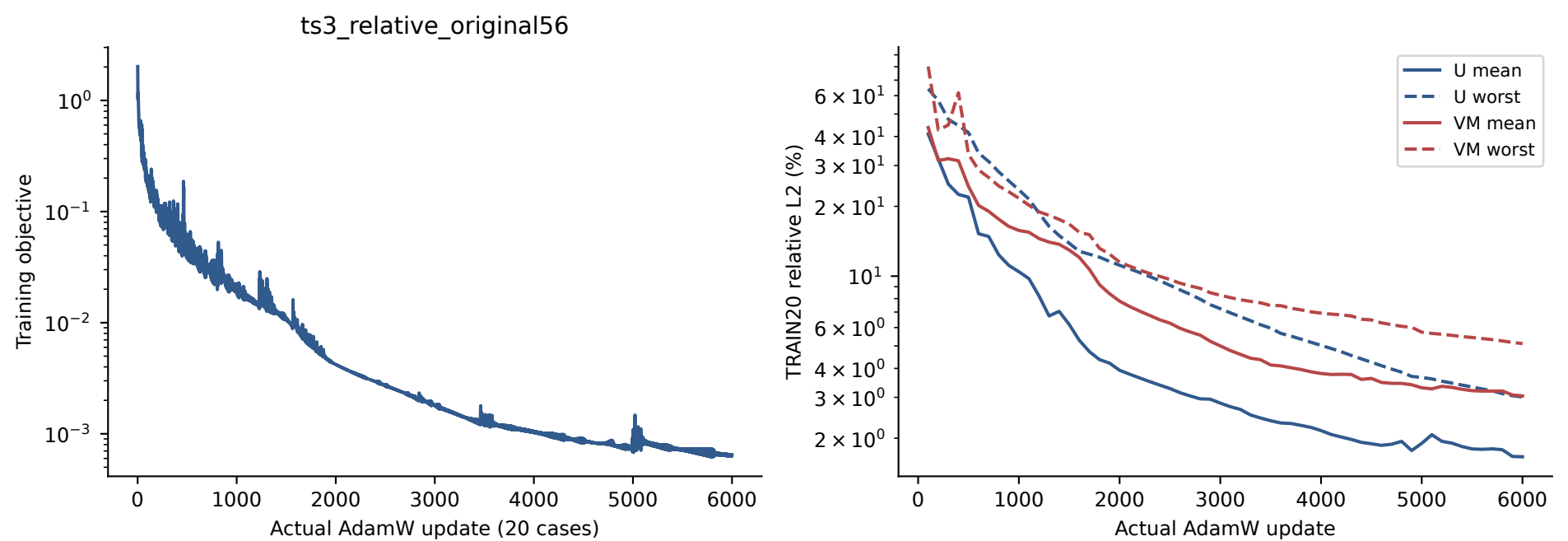

ts3_relative_original56
Training objective
10^{0}
10^{-1}
10^{-2}
10^{-3}
0
1000
2000
3000
4000
5000
6000
Actual AdamW update (20 cases)
TRAIN20 relative L2 (%)
6×10^{1}
4×10^{1}
3×10^{1}
2×10^{1}
10^{1}
6×10^{0}
4×10^{0}
3×10^{0}
2×10^{0}
U mean
U worst
VM mean
VM worst
Actual AdamW update