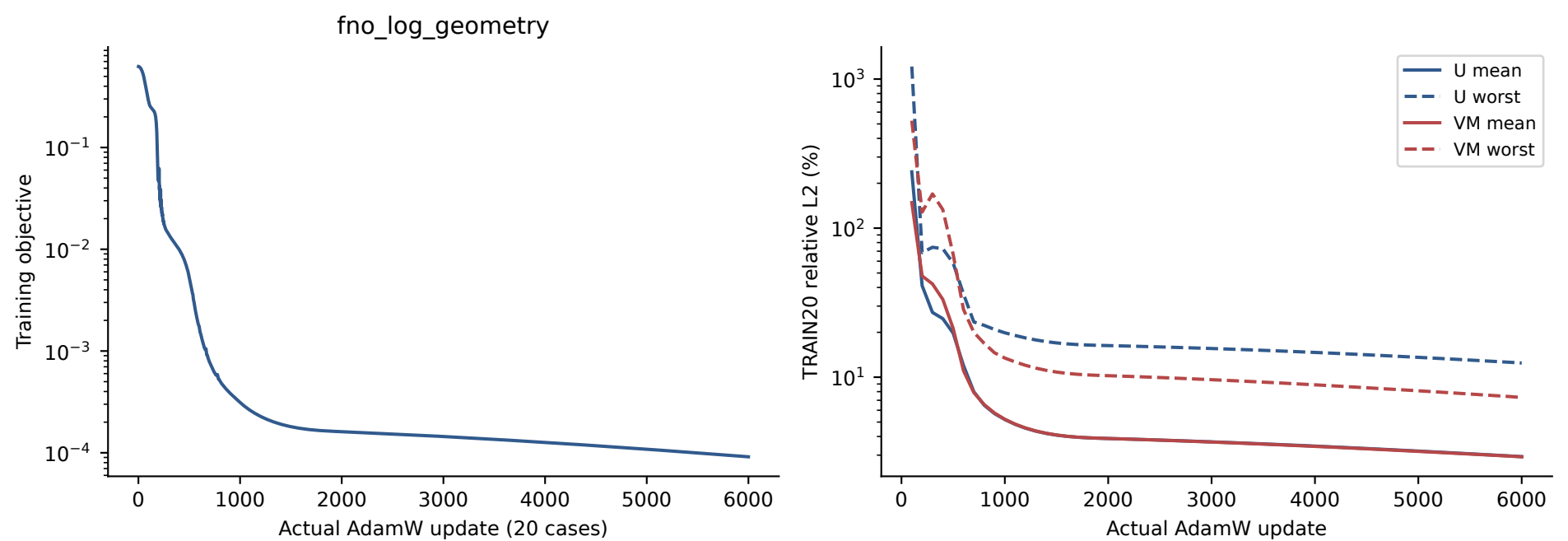

fno_log_geometry
Training objective
$10^{-1}$
$10^{-2}$
$10^{-3}$
$10^{-4}$
0
1000
2000
3000
4000
5000
6000
Actual AdamW update (20 cases)
TRAIN20 relative L2 (%)
$10^{3}$
$10^{2}$
$10^{1}$
U mean
U worst
VM mean
VM worst
0
1000
2000
3000
4000
5000
6000
Actual AdamW update