

AWC · four 20M workers · no clipping · learning-rate response
Mean $\pm$ sample SD across seeds 42, 43, 44 · shared loss scale

