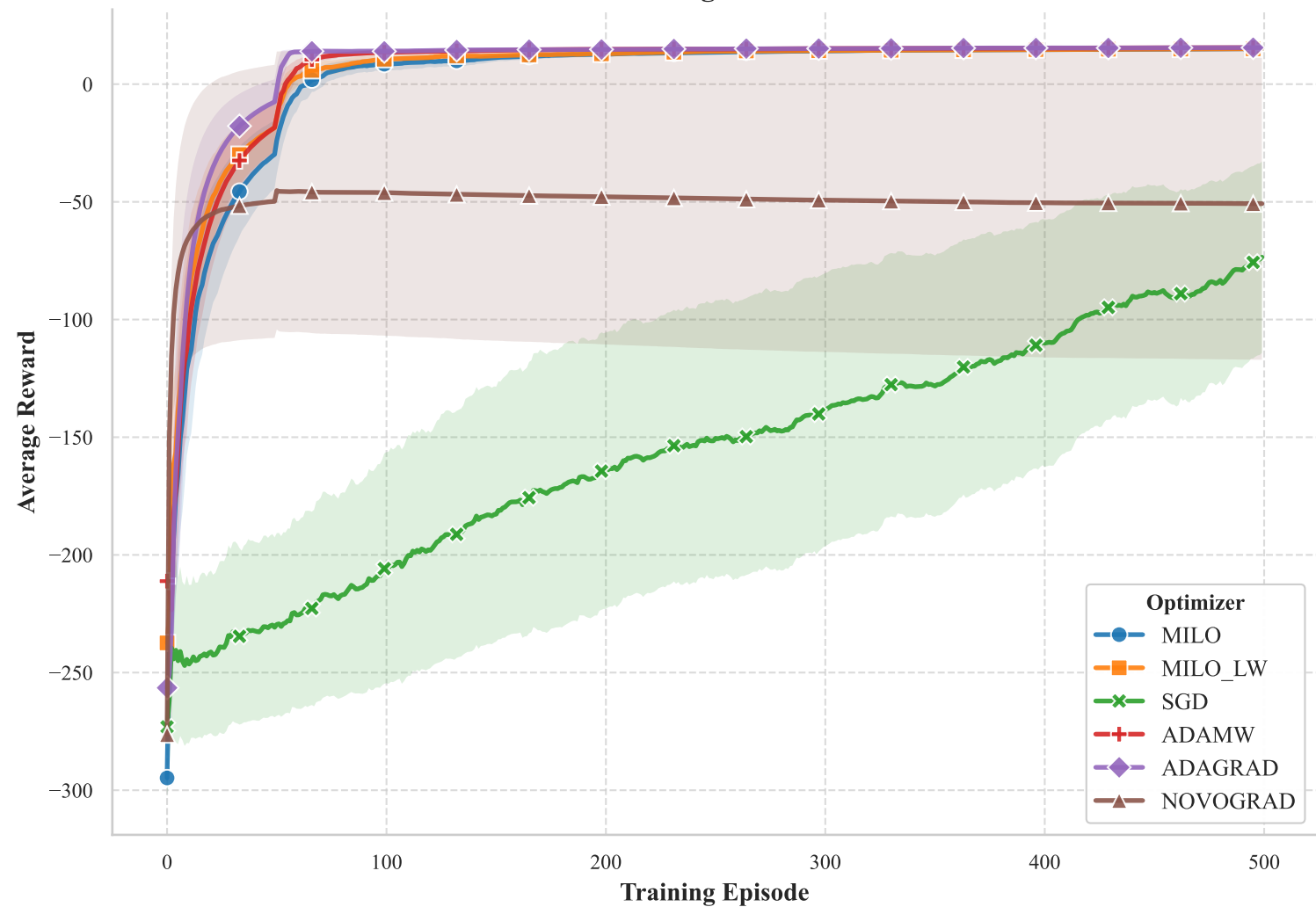


Training Reward



Shaded areas represent standard error of the mean across runs