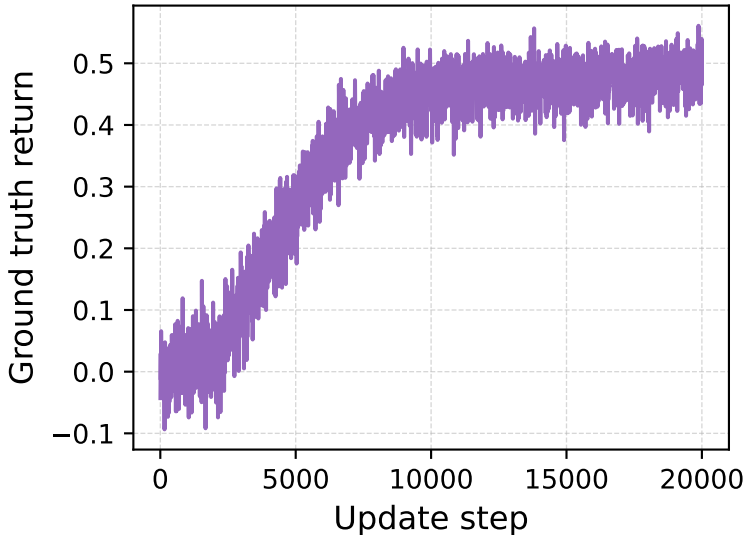


Learning curves at 5.0% oversight



Gradient-routed MoE