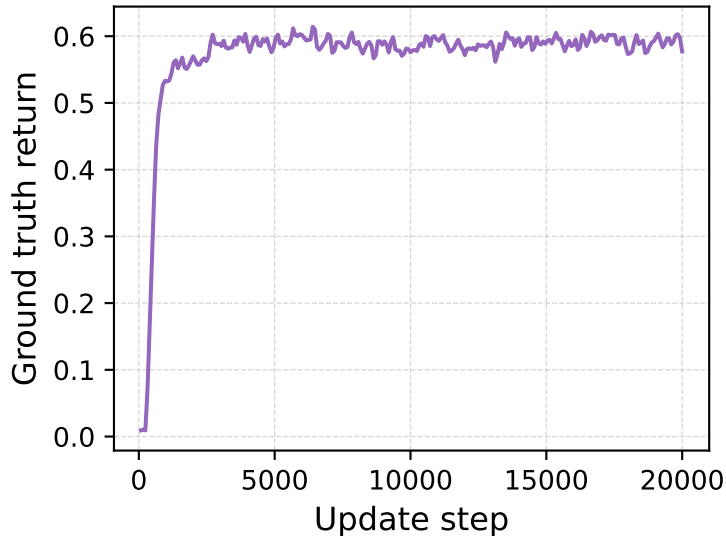


Learning curves at 80.0% oversight



Gradient-routed MoE