

Mean Norm Trajectory Comparison: SFT vs GRPO

