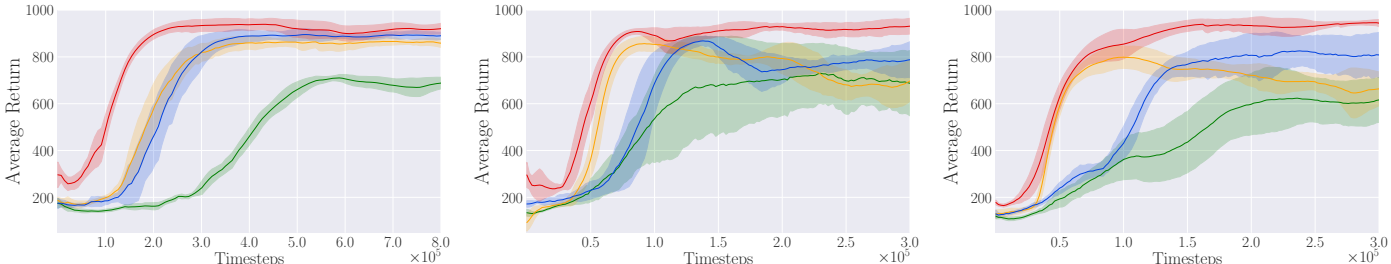

Fig. 5. Benchmarks of four RL frameworks trained with (a) PPO, (b) TD3, and (c) SAC RL algorithms. Each plot depicts the learning curves for Mono-MLP (green), Mono-**EMLP** (blue), Mod-MLP (orange), and Mod-**EMLP** (red).