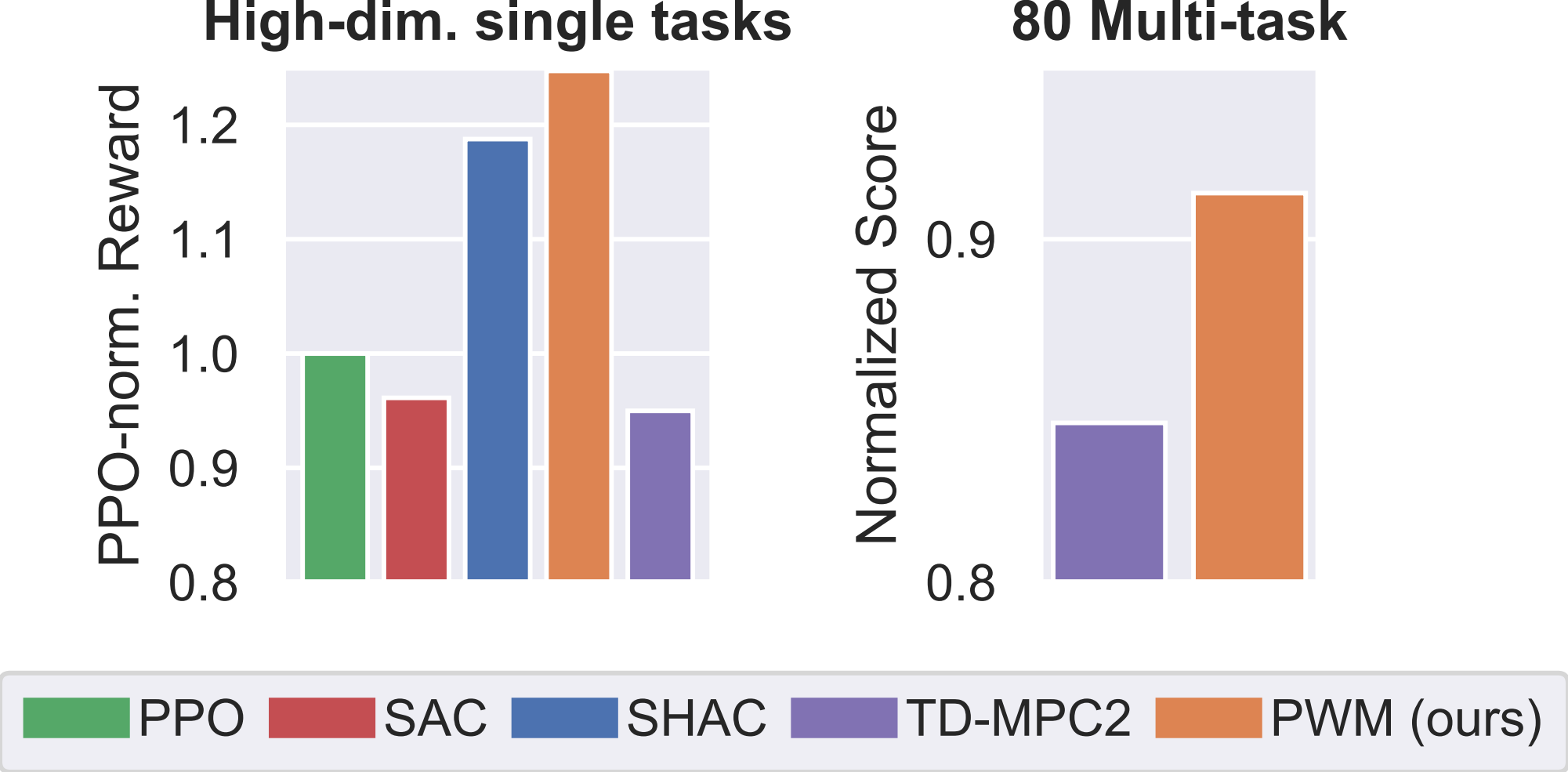
High-dim. single tasks
PPO-norm. Reward
1.2
1.1
1.0
0.9
0.8
80 Multi-task
Normalized Score
0.9
0.8
PPO
SAC
SHAC
TD-MPC2
PWM (ours)