Mean Reward of 20 trajectories ± standard deviation

Legend:
- SAC
- $\alpha=0$
- $\alpha=1$
- $\alpha=2$
- $\alpha=3$
- $\alpha=4$
- $\alpha=5$

Y-axis: Mean Reward

X-axis: Number of steps (1e6)