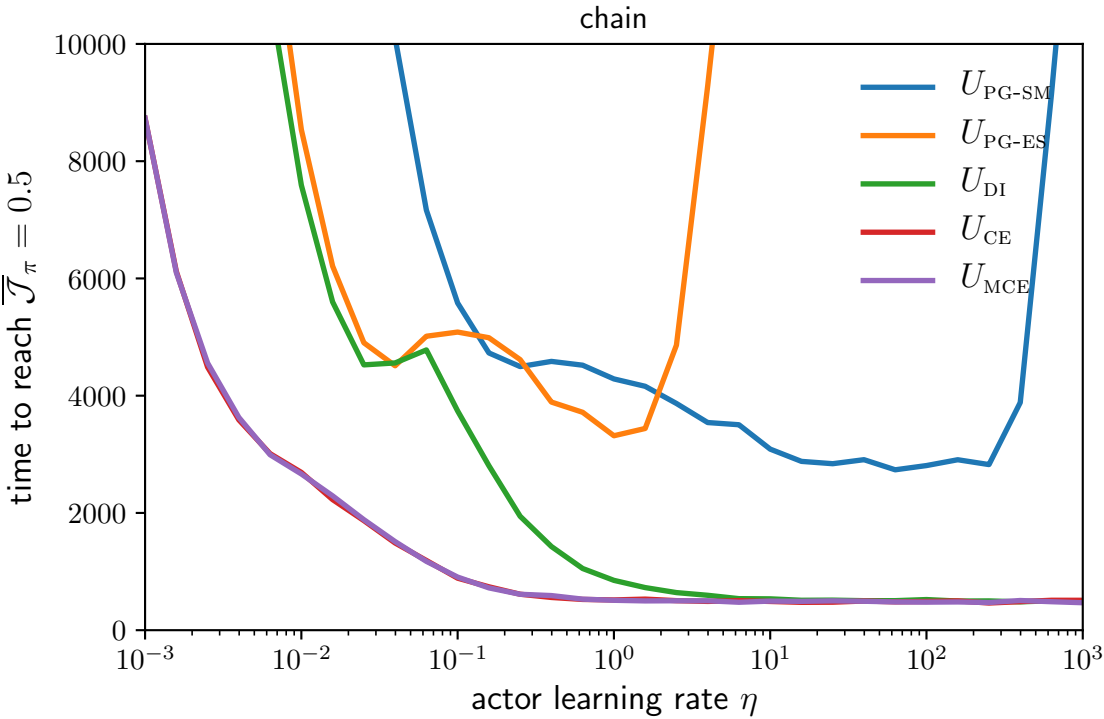
chain
time to reach $\overline{\mathcal{J}}_\pi = 0.5$
actor learning rate η
$U_{\text{PG-SM}}$
$U_{\text{PG-ES}}$
U_{DI}
U_{CE}
U_{MCE}
10000
8000
6000
4000
2000
0
10^{-3}
10^{-2}
10^{-1}
10^{0}
10^{1}
10^{2}
10^{3}