

#(inner updates): 1

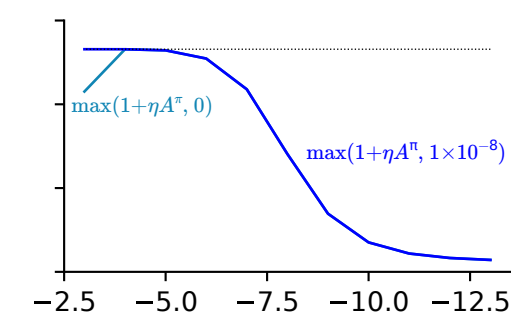
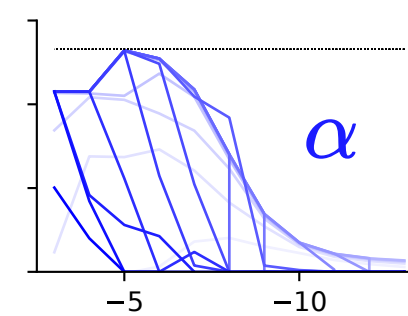
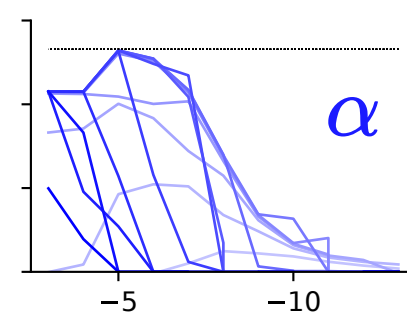
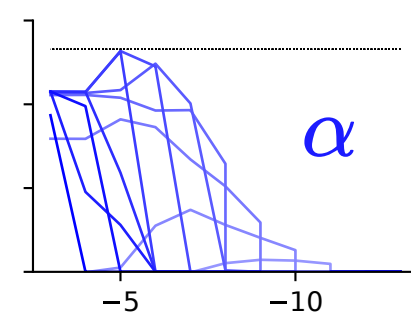
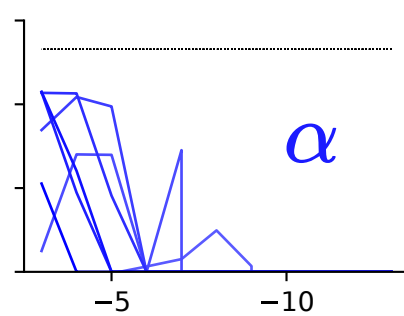
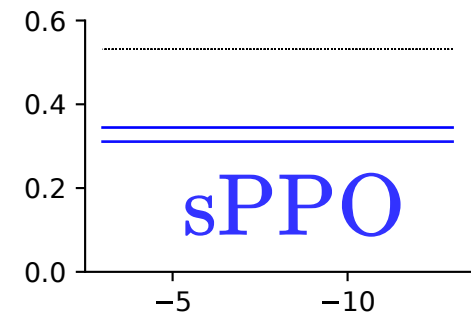
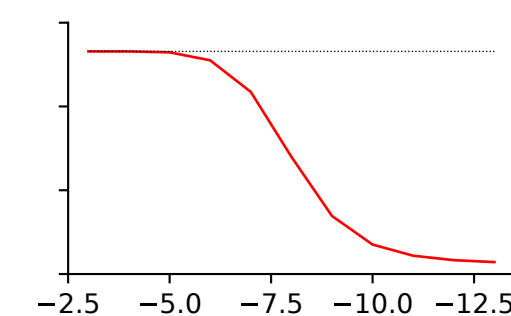
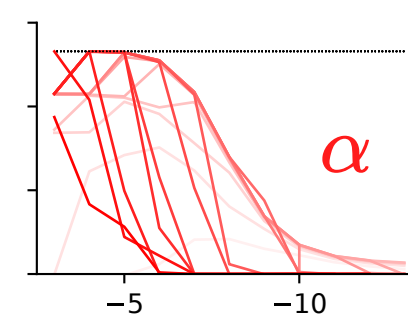
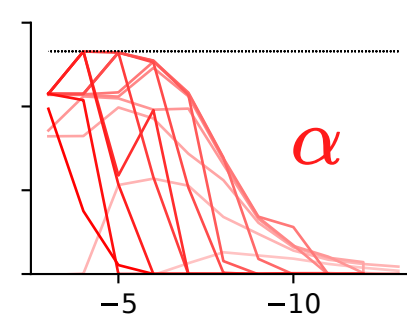
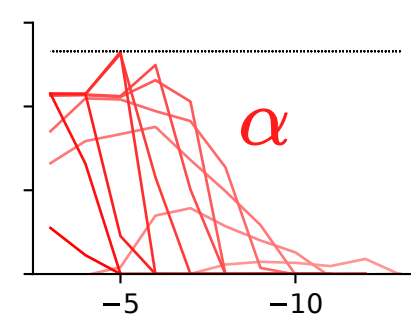
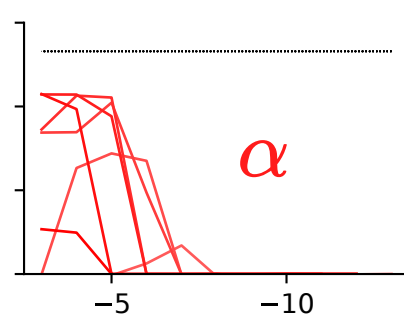
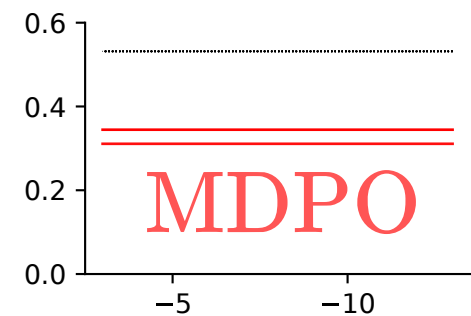
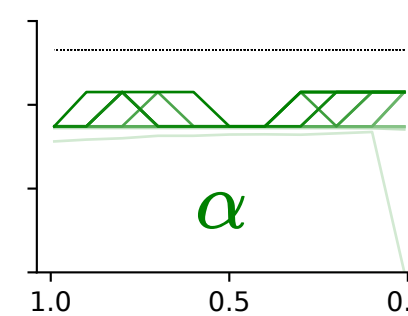
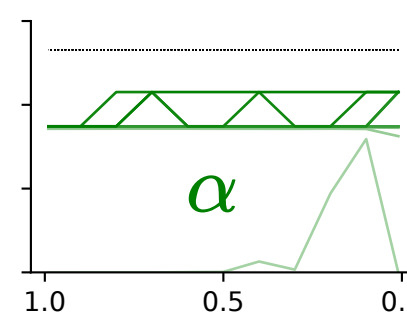
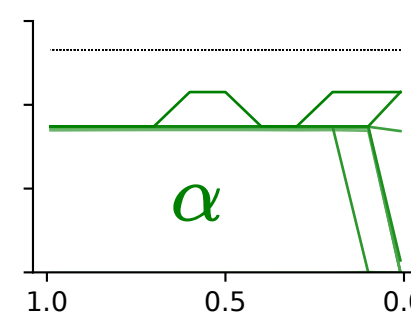
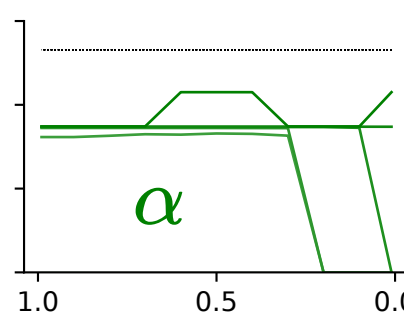
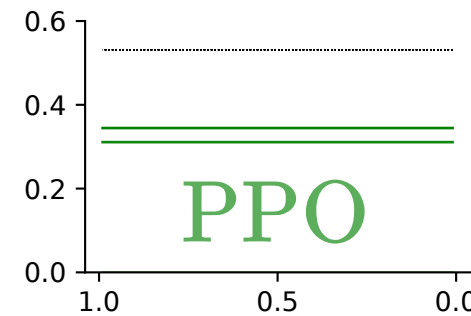
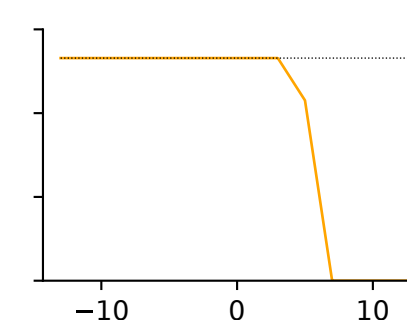
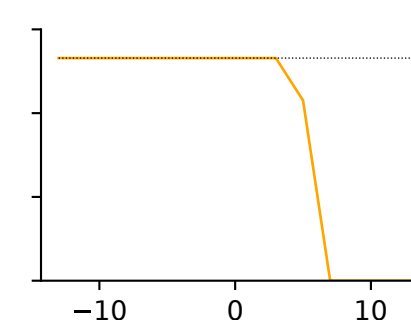
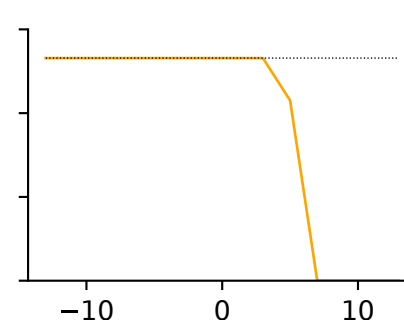
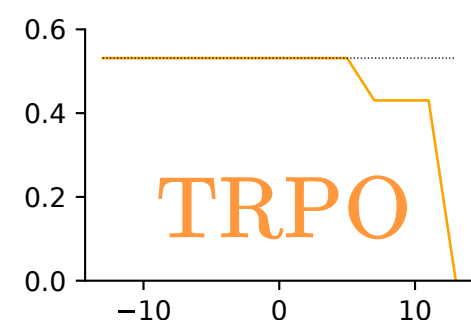
10

100

1000

10000

Analytical

 x axis labelFinal Value of \mathcal{J} (PG objective) $\log_2(\eta)$ Increasing
regularization
strength ($1/\eta$) $\log_2(\eta)$ Increasing
regularization
strength ($1/\eta$) ϵ Increasing
epsilon constraint

Note: TRPO's performance already saturated at 10 number of inner updates: it has almost the same performance at 10, 100, and 1000 number of inner updates. Therefore, we didn't run it for 10000 updates; it should be very similar to parameter sensitivity plots for 10, 100, and 1000 number of inner updates.

 $\log_2(\delta)$ Increasing KL
constraint strength