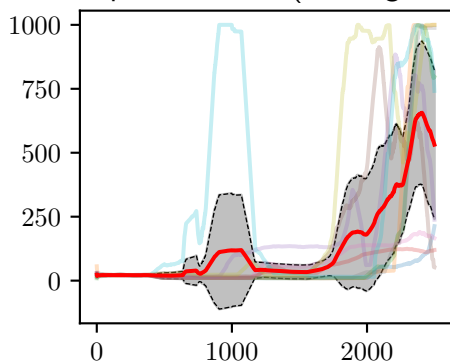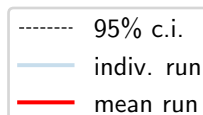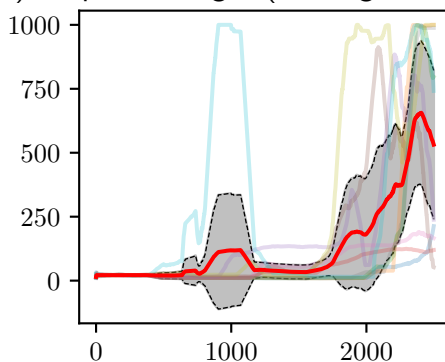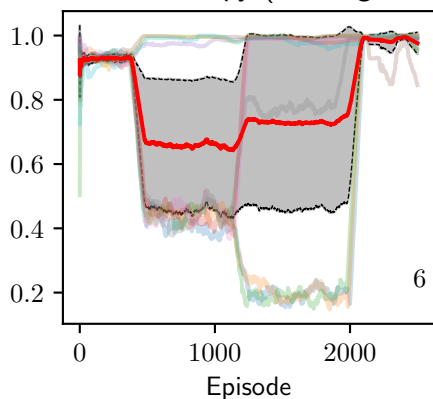MC training progress on cart-pole task

100-episode return (running mean)   Episode length (running mean)

Behavior entropy (running mean)   Sample loss (RMSE) (running mean)

Legend:
- 95% c.i.
- indiv. run
- mean run