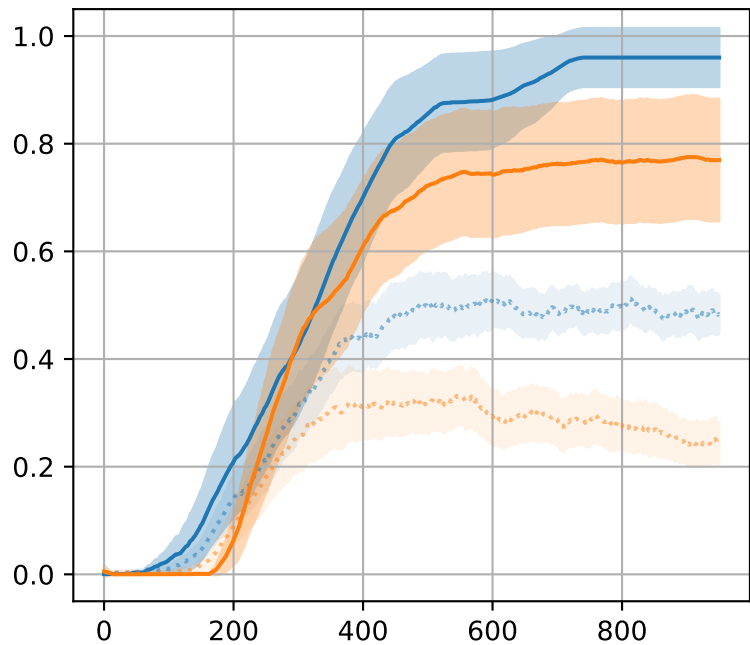
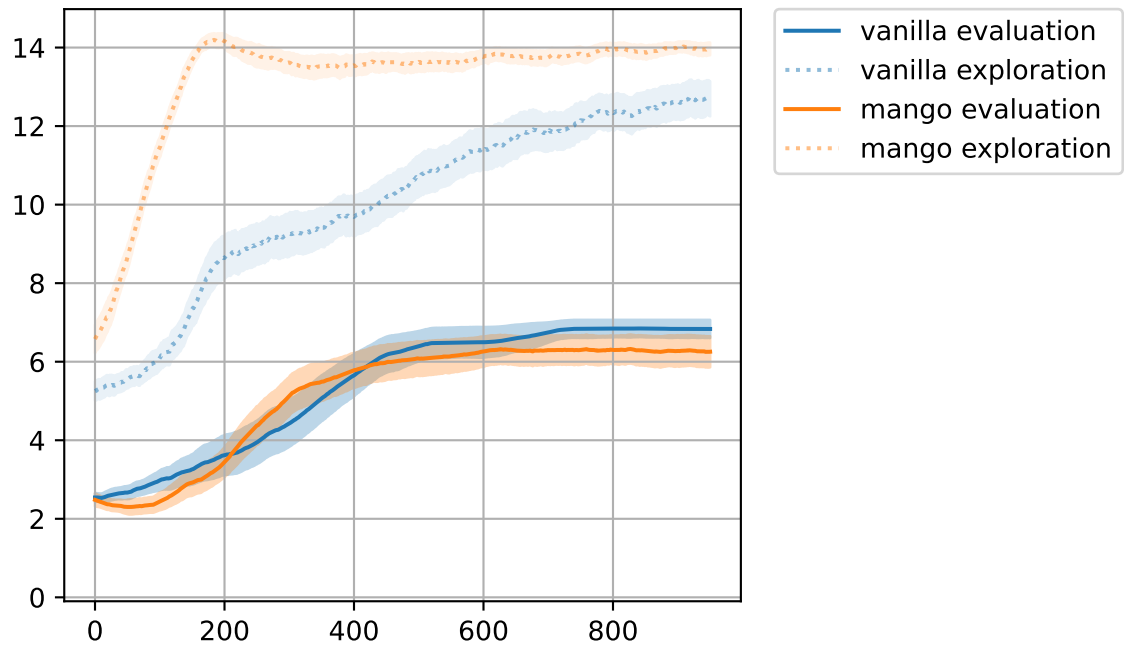


average across 50 runs

reward



episode length



- vanilla evaluation
- vanilla exploration
- mango evaluation
- mango exploration