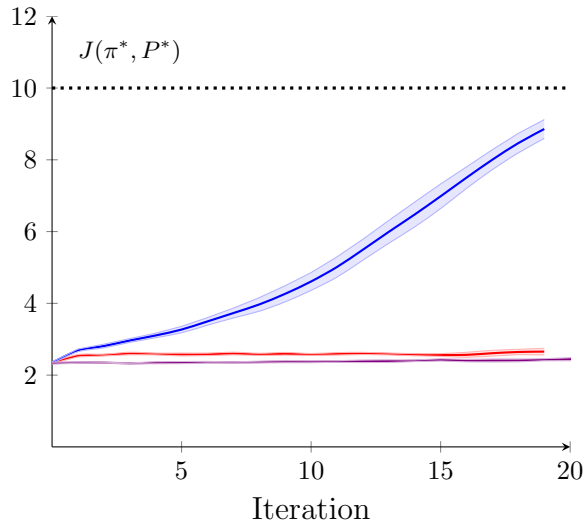
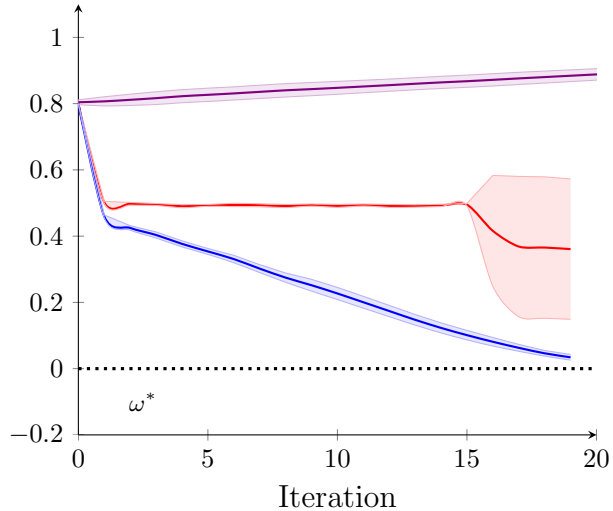


Average Reward

 \mathfrak{z} 

- G(PO)MDP
- REMPS $\epsilon = 1e-4$
- REMPS $\epsilon = 1e-2$