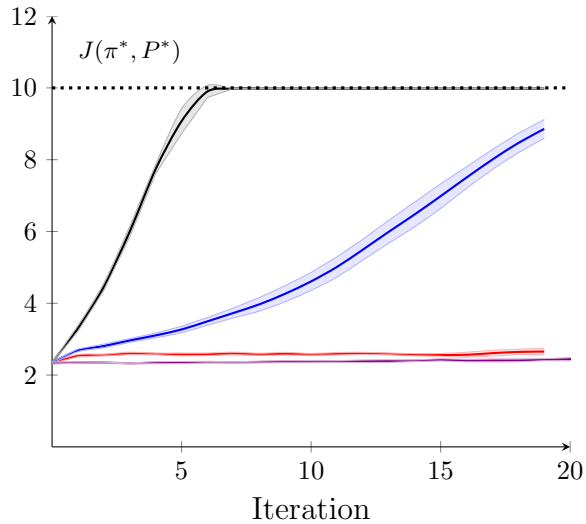
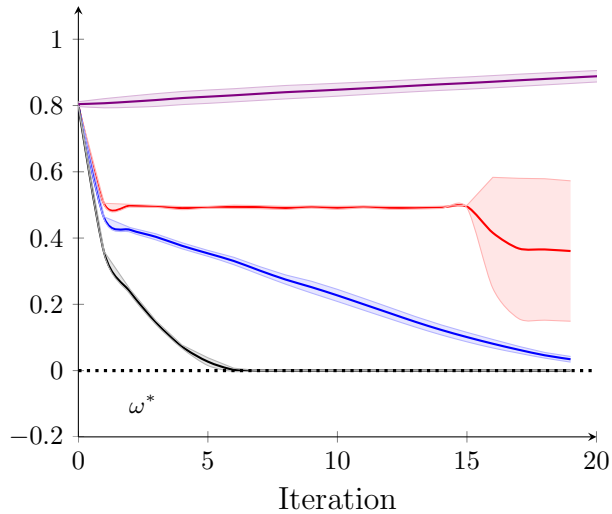


Average Reward

 ω 

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