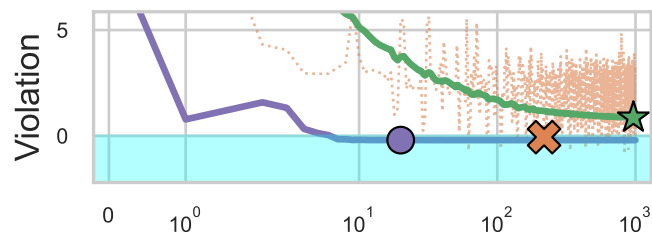


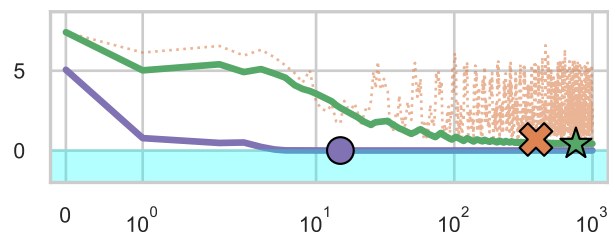


(a) RCMDP ($|\mathcal{U}| = 5$)



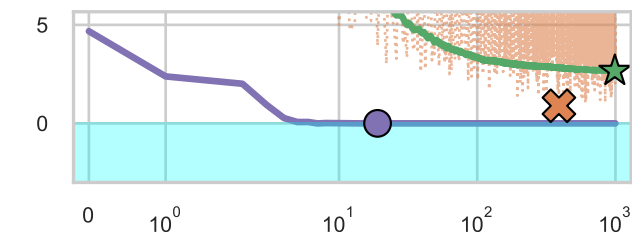
Outer Iteration k

(b) RCMDP (KL)



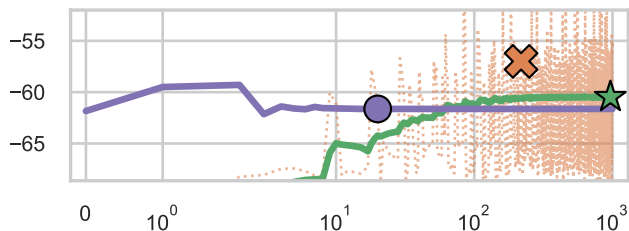
Outer Iteration k

(c) CMDP ($|\mathcal{U}| = 1$)

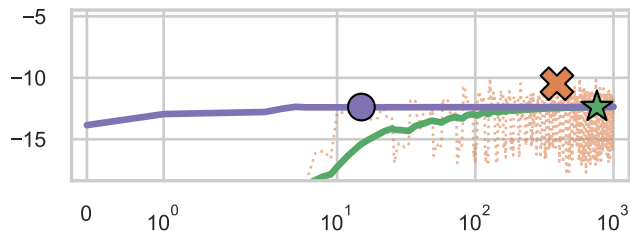


Outer Iteration k

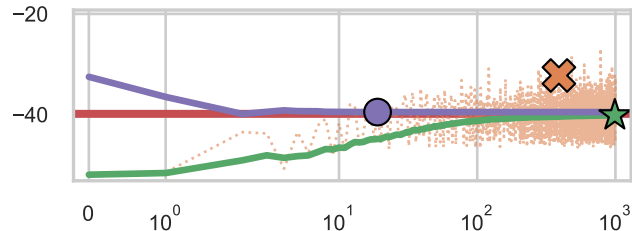
Objective Return
(Smaller is better)



Outer Iteration k

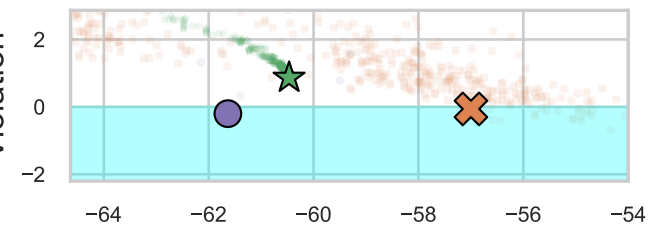


Outer Iteration k

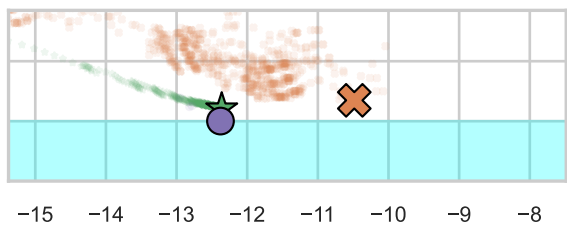


Outer Iteration k

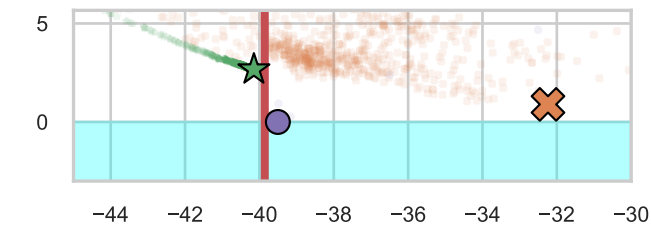
Violation



Objective Return
(smaller is better)



Objective Return
(smaller is better)



Objective Return
(smaller is better)