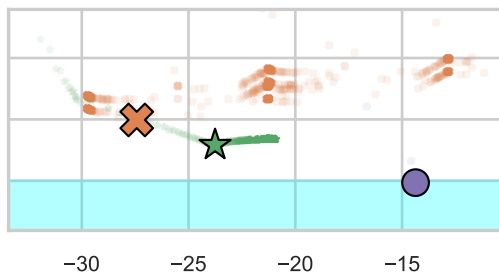


feasible region LF LF- π^k -avg **EpiRC-PGS (Ours)**

LF- $d_P^{\pi^k}$ -avg (only for $|\mathcal{U}| = 1$)

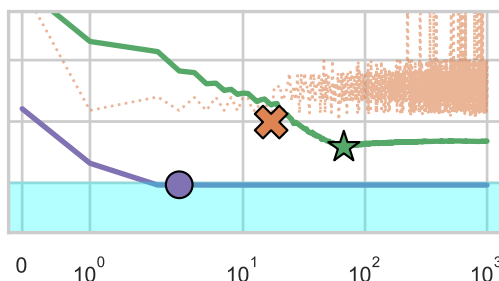
RCMDP (KL)

Constraint Violation
(Blue area is feasible)



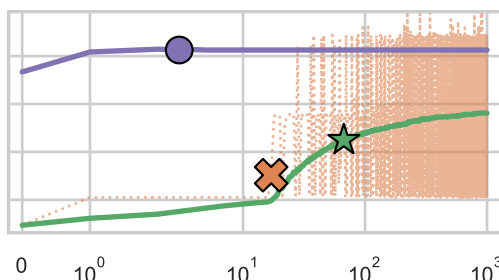
Objective Return
(smaller is better)

Constraint Violation
(Blue area is feasible)



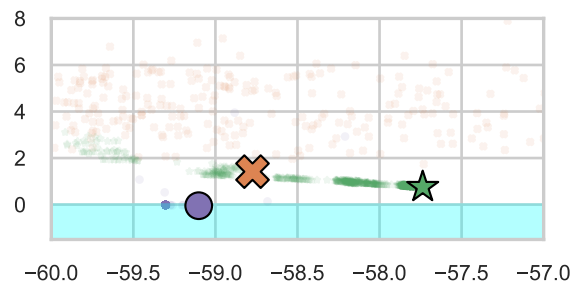
Outer Iteration k

Objective Return
(Smaller is better)

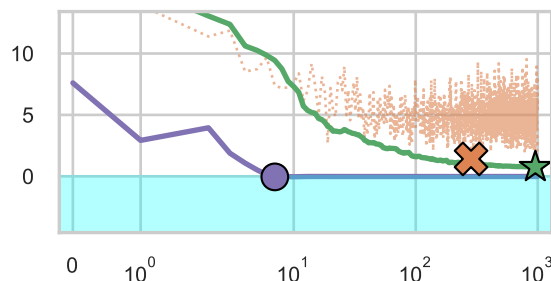


Outer Iteration k

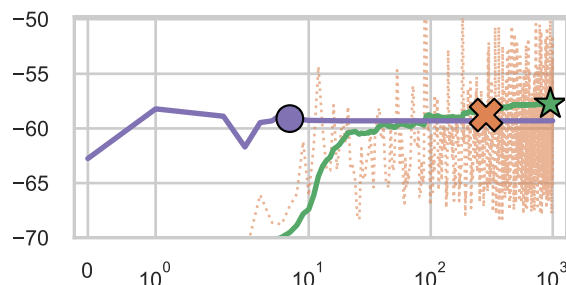
RCMDP ($|\mathcal{U}| = 5$)



Objective Return
(smaller is better)

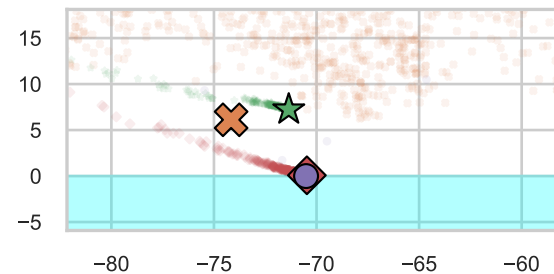


Outer Iteration k

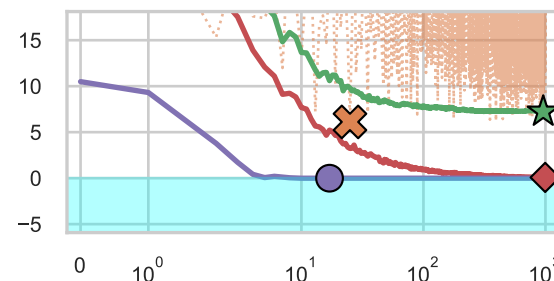


Outer Iteration k

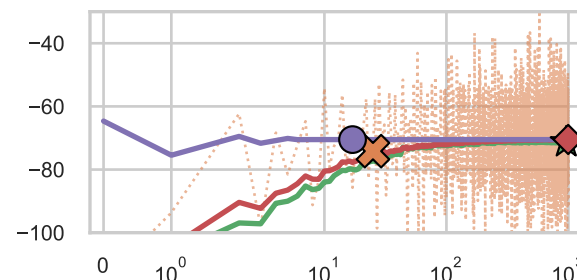
CMDP ($|\mathcal{U}| = 1$)



Objective Return
(smaller is better)



Outer Iteration k



Outer Iteration k