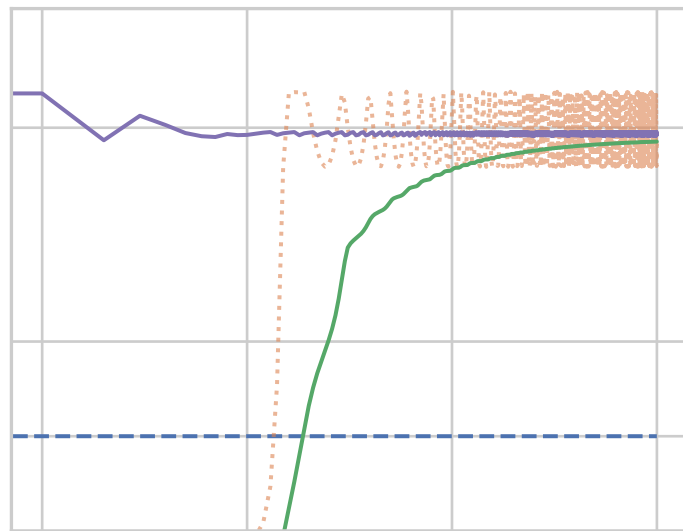


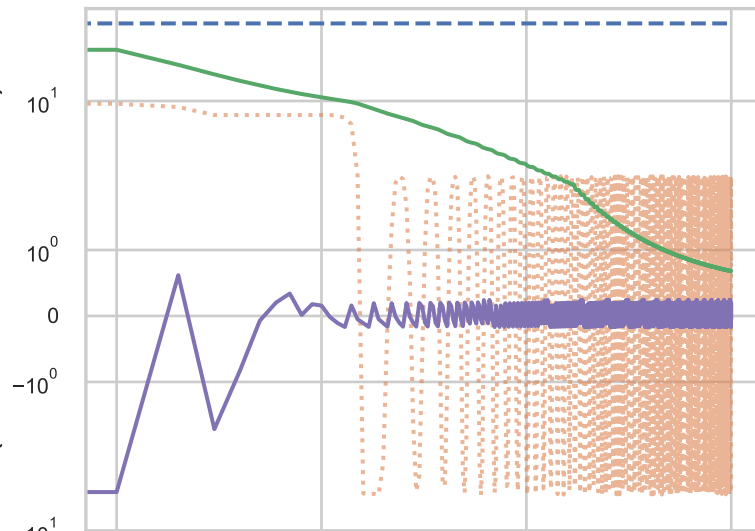
--- Uniform policy ( $\pi_{\text{unif}}$ )    ..... LF-PGS    — LF-PGS-avg    — **EpiRC-PGS (Ours)**

Relative Objective Return  
(Smaller is better)



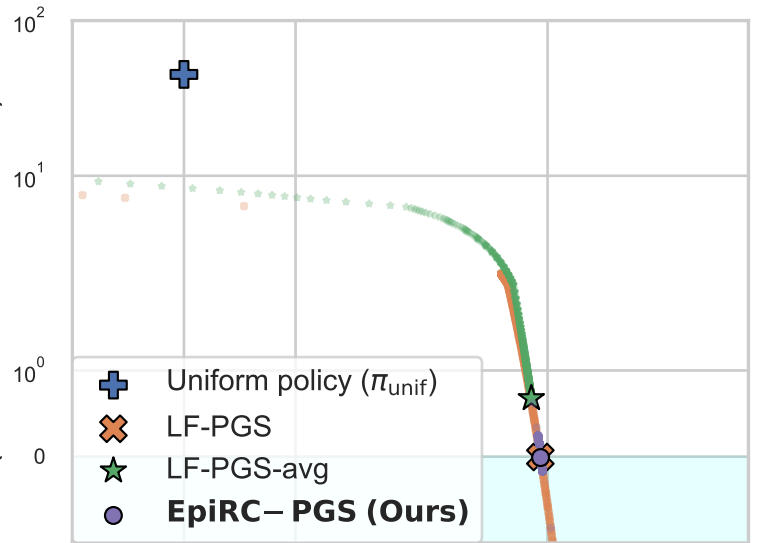
Outer Iteration  $k$

Constraint Violation  
(Blue area is feasible)



Outer Iteration  $k$

Constraint Violation  
(Blue area is feasible)



+ Uniform policy ( $\pi_{\text{unif}}$ )  
× LF-PGS  
★ LF-PGS-avg  
● **EpiRC-PGS (Ours)**

Relative Objective Return  
(Smaller is better)