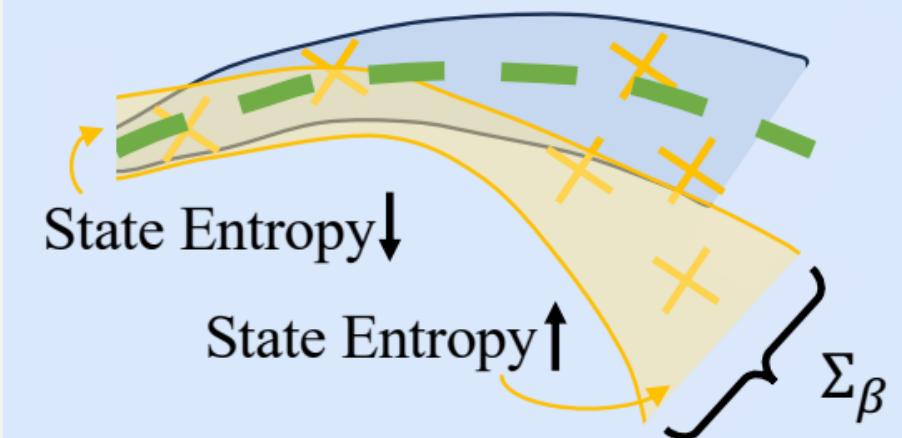


X Suboptimal Intervention Data  
- Optimal Policy    □  $\pi$     □  $\beta$



$$\max_{\lambda} \inf_{\pi} -\mathcal{J}(\pi) + \lambda^T [D(\pi, \beta) - \Sigma_\beta]$$