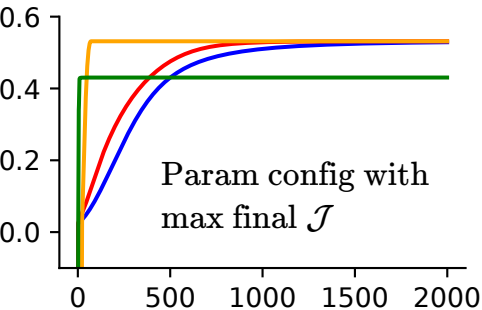
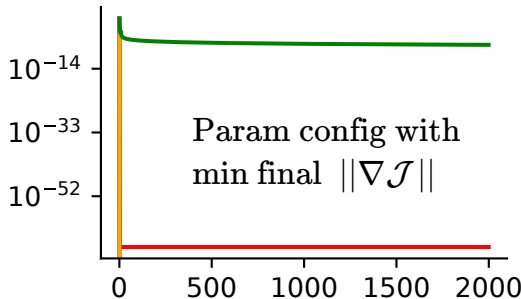
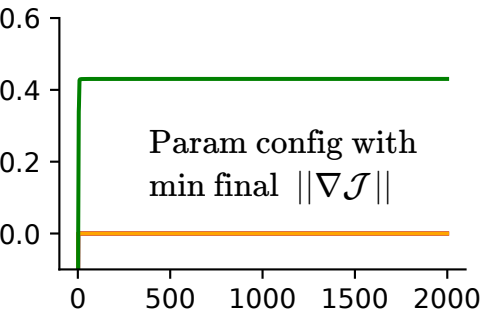
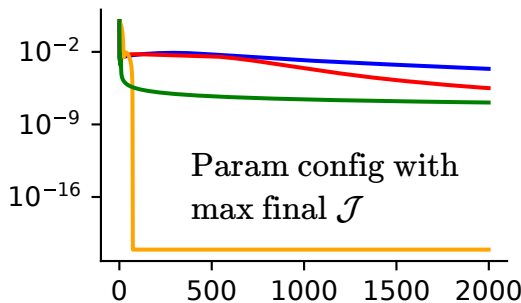


\mathcal{J} (PG objective) $\|\nabla \mathcal{J}\|$ (policy gradient)

of iterations