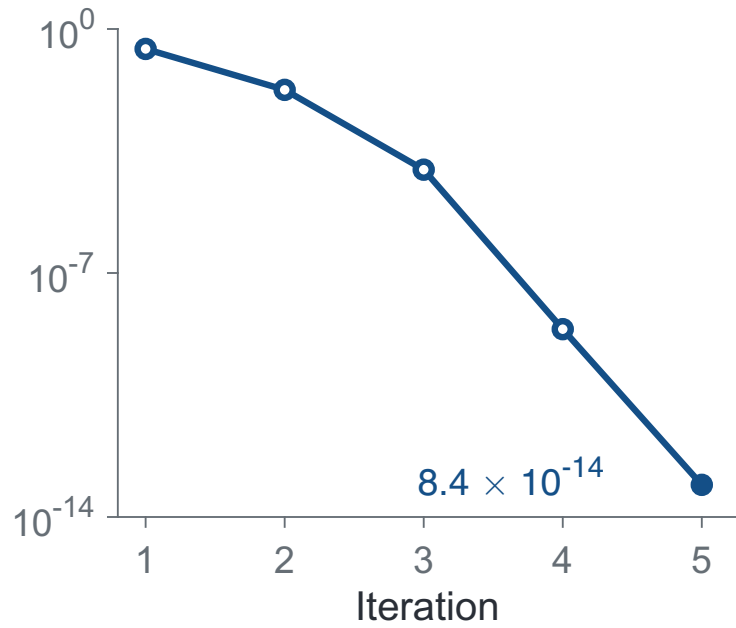
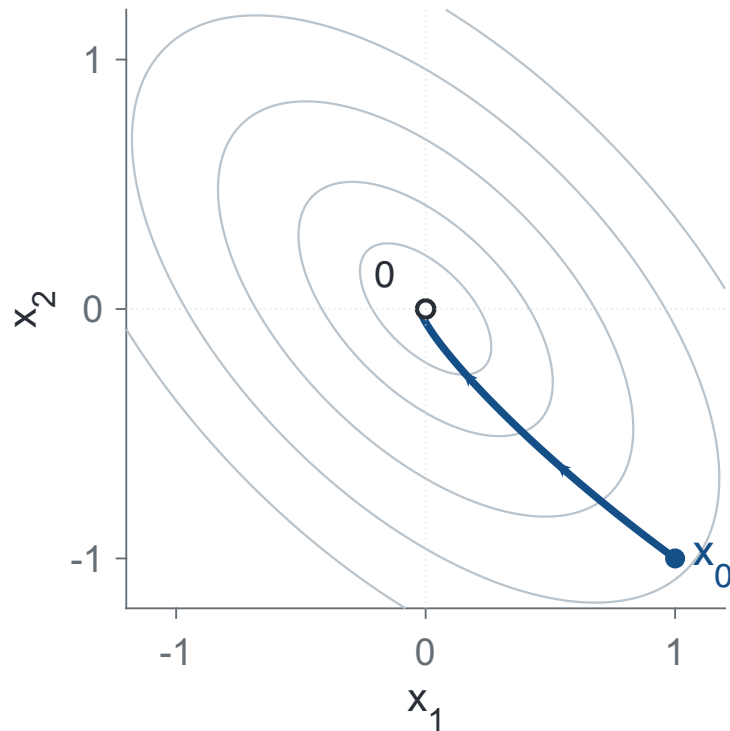


a Policy learning



b Closed-loop trajectory



c Value decomposition

