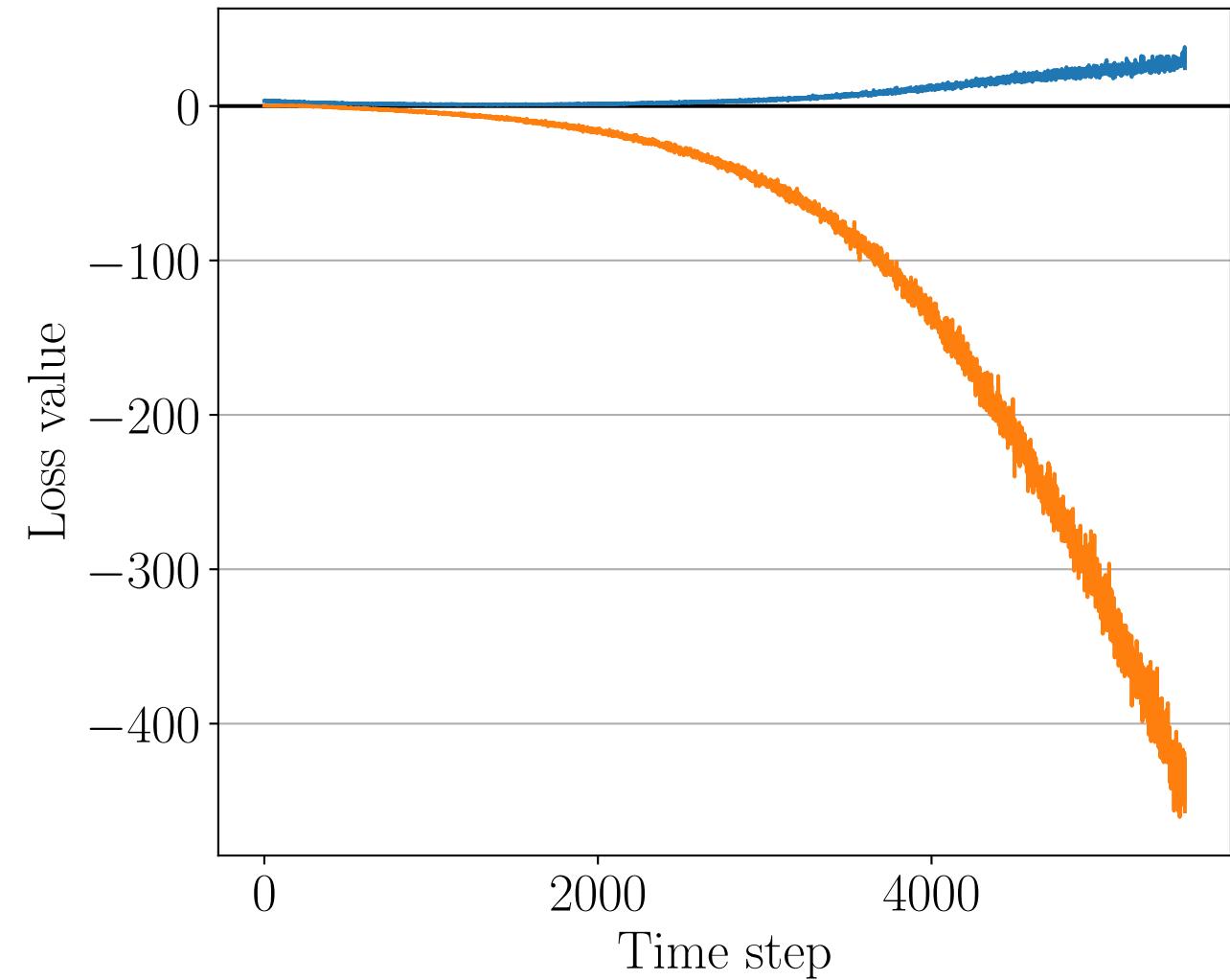
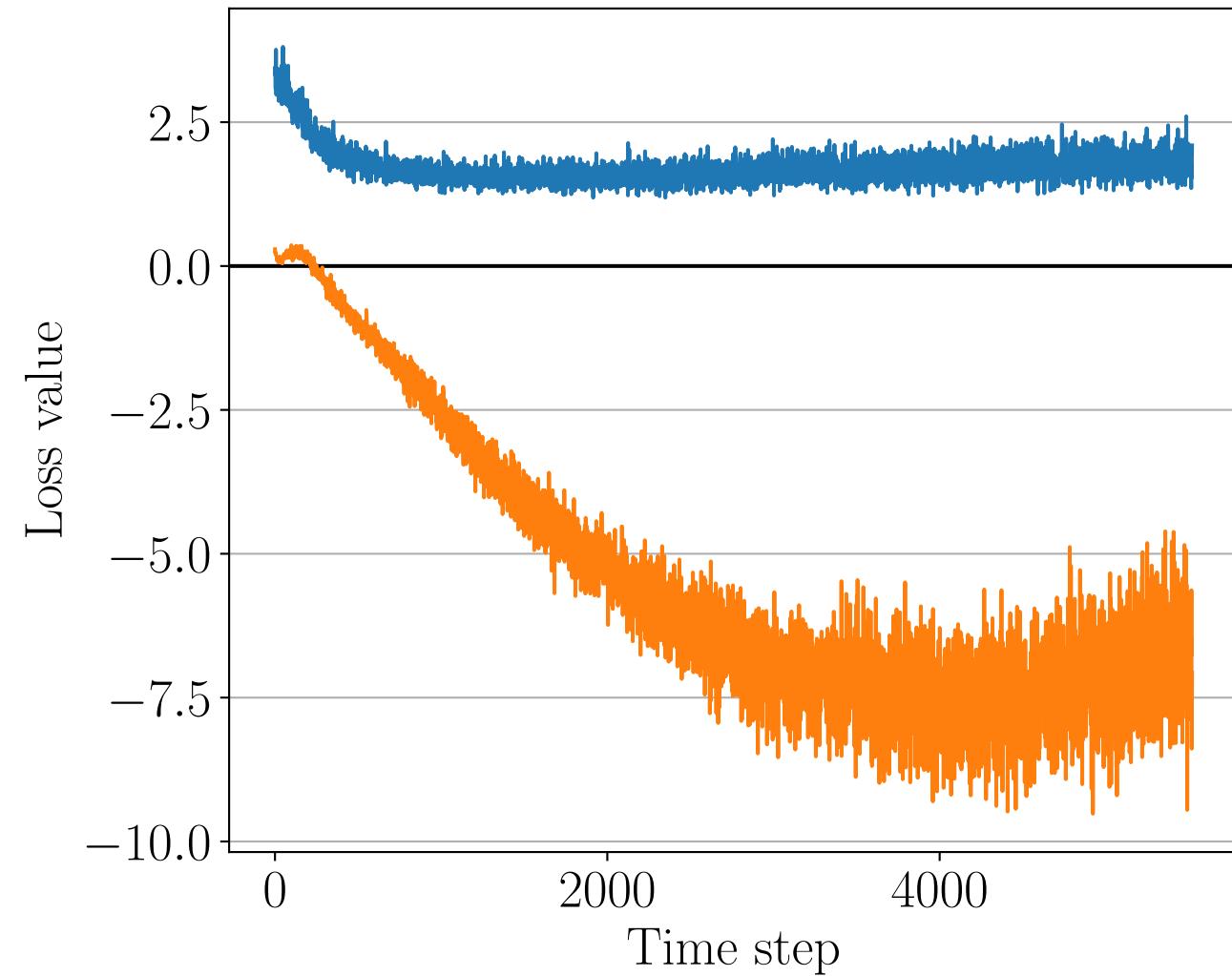


a) $\alpha_i = 0.15$



b) $\alpha_i = 1e^{-4}$



— Q-Network loss
— Policy loss