(a) Training Reward
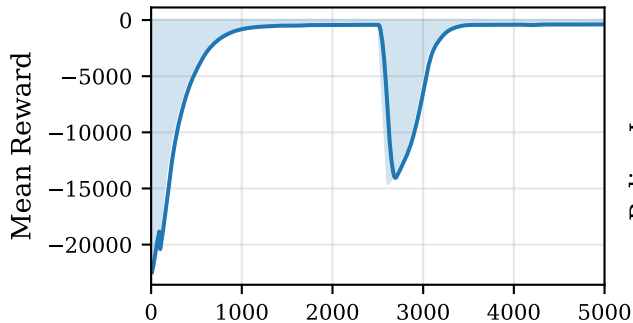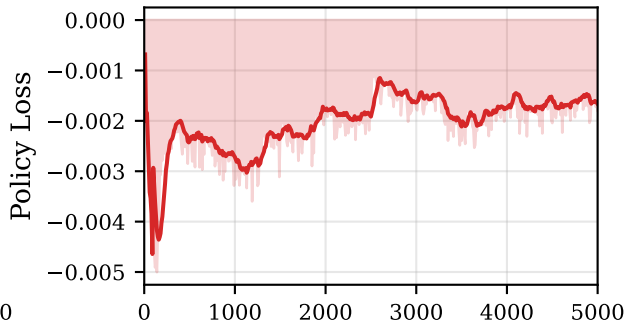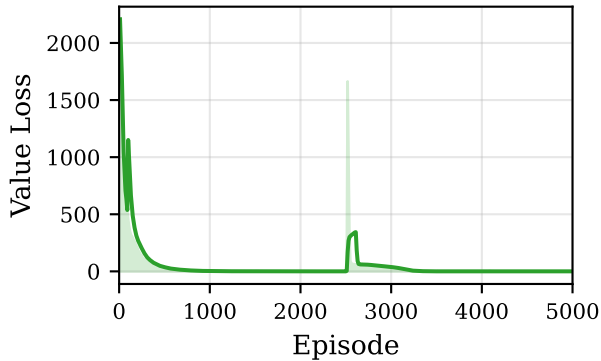
(b) Policy Loss

(c) Value Loss

(d) Policy Entropy