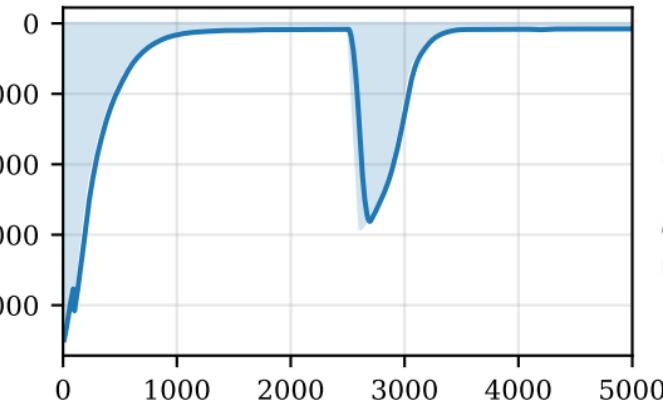


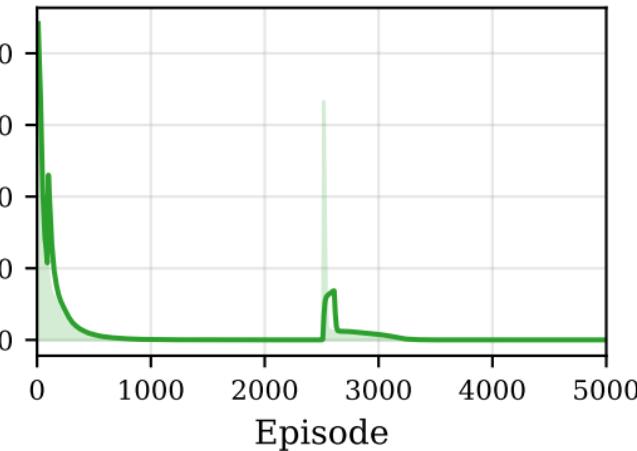
(a) Training Reward

Mean Reward



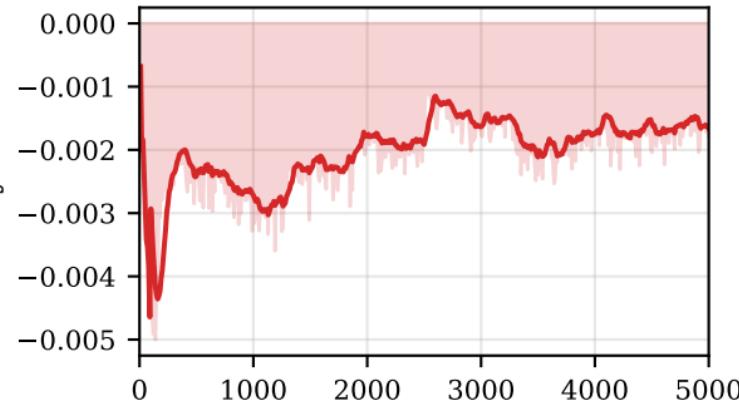
Episode
(c) Value Loss

Value Loss



(b) Policy Loss

Policy Loss



Episode
(d) Policy Entropy

Entropy

