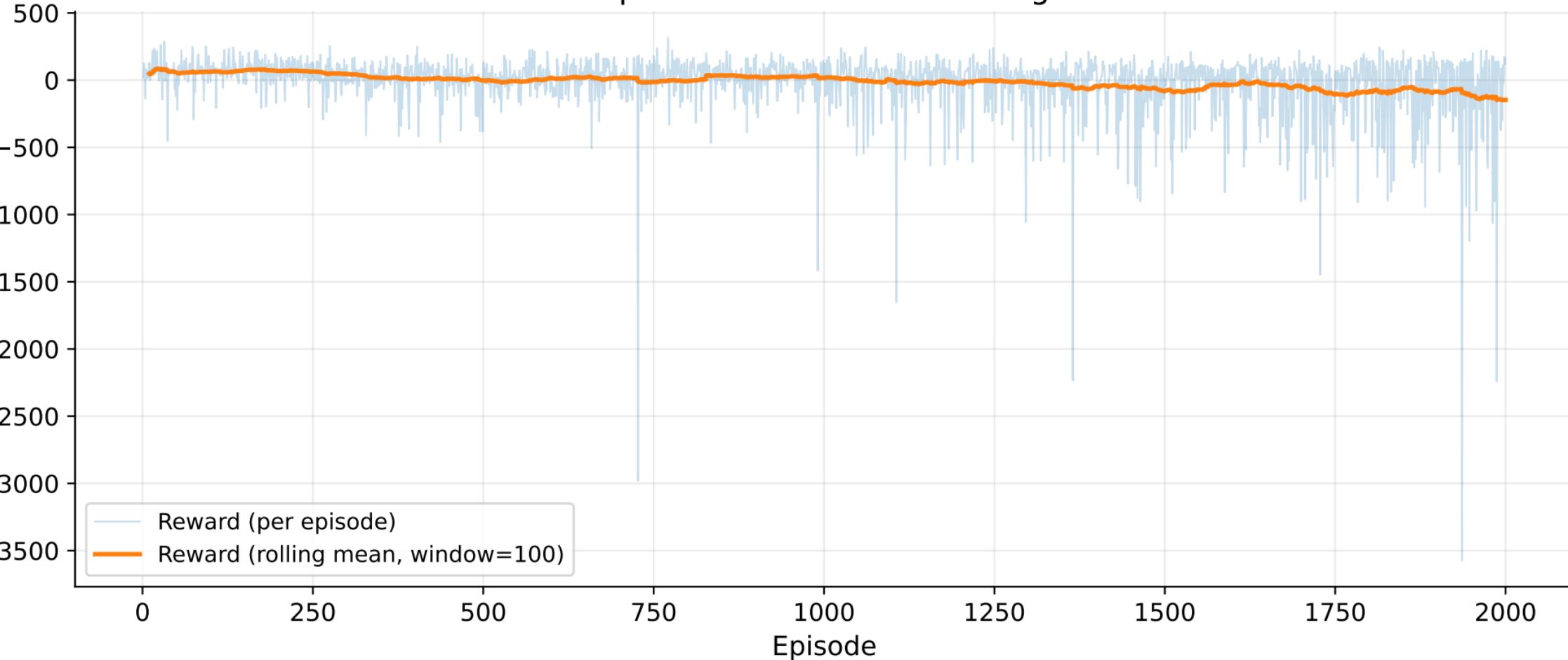
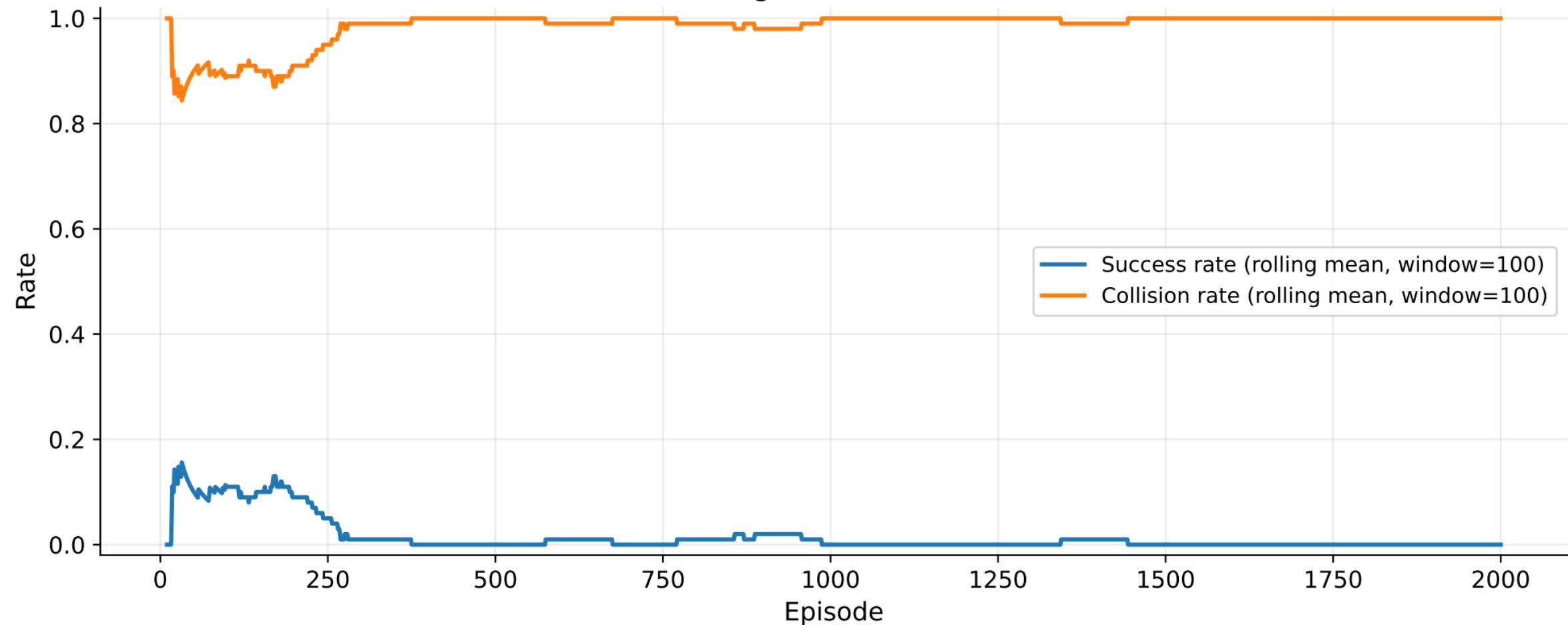


Episode reward over training

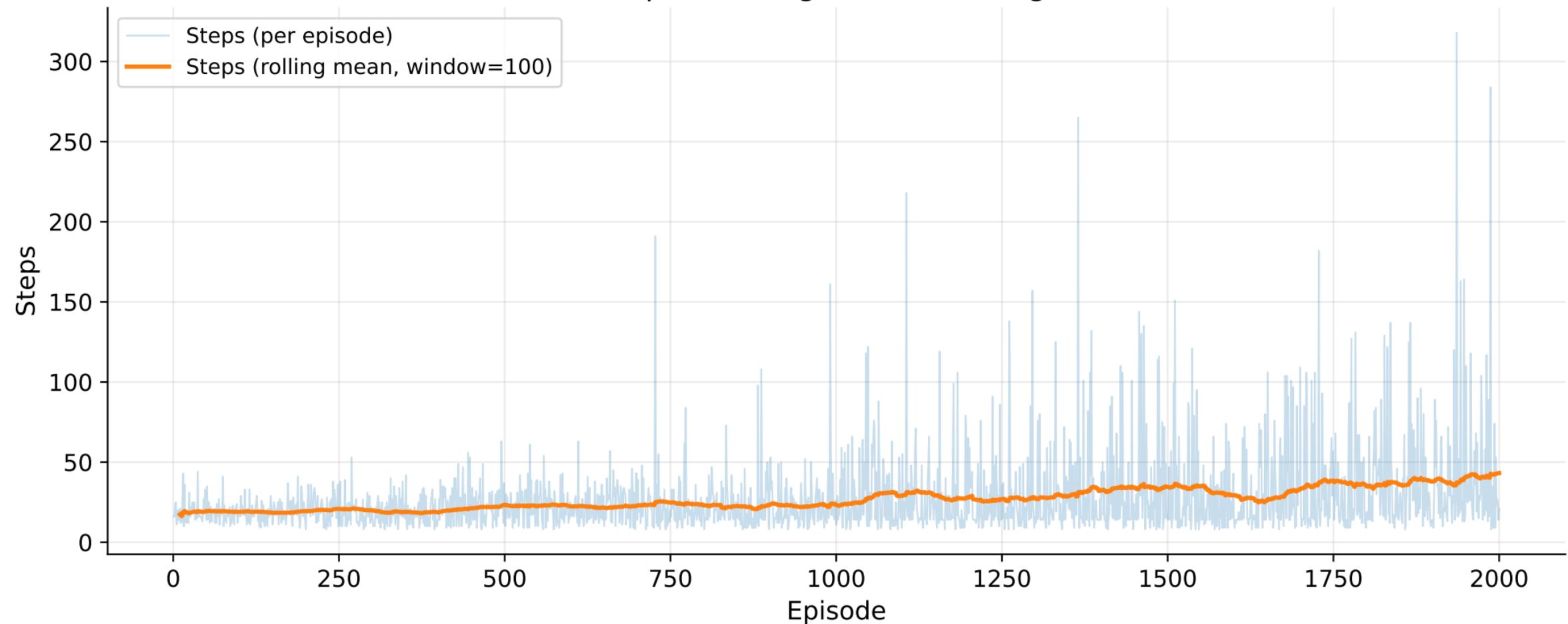
Reward



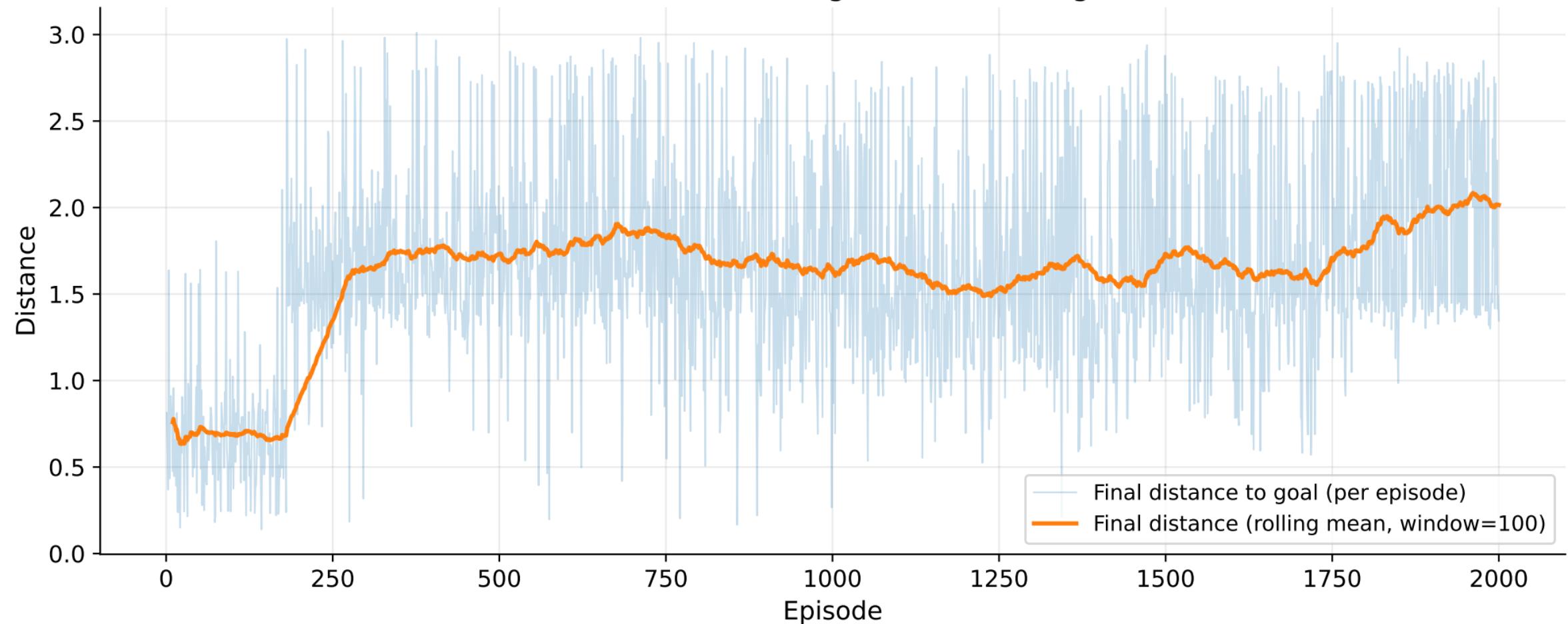
Rolling outcome rates



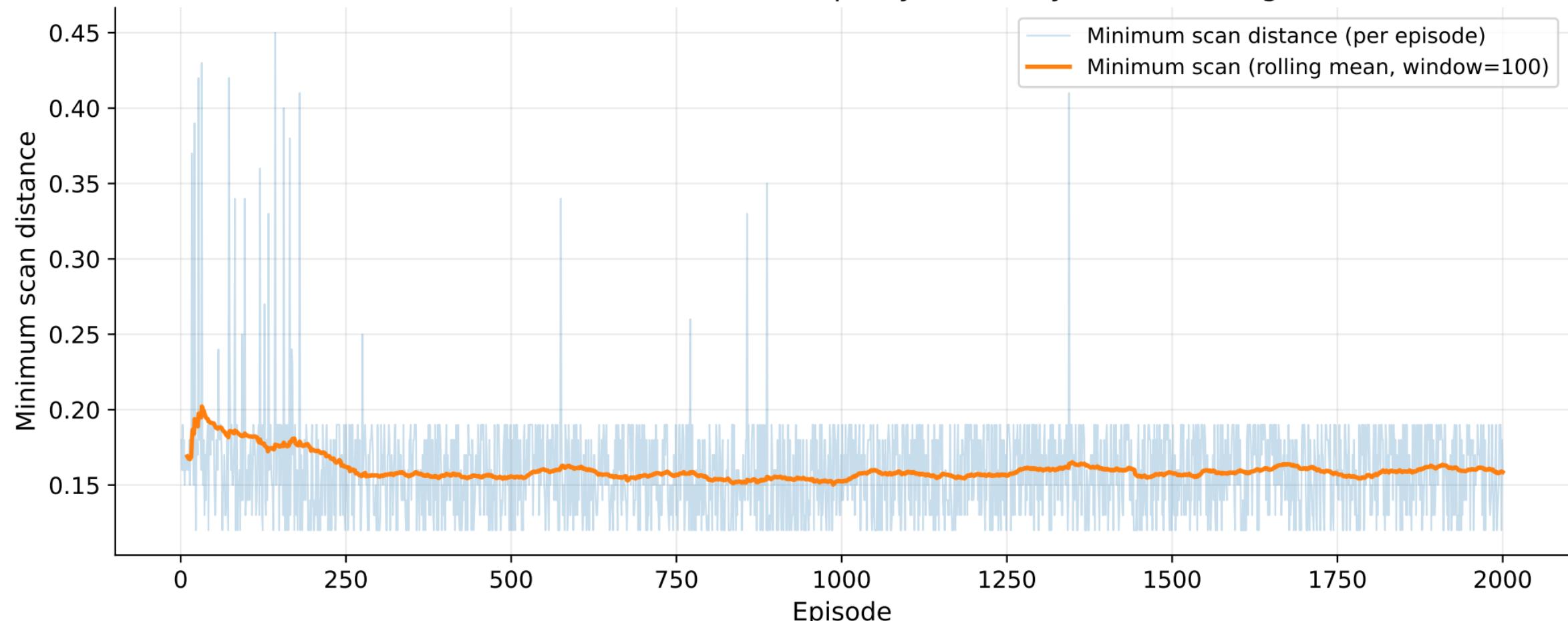
Episode length over training



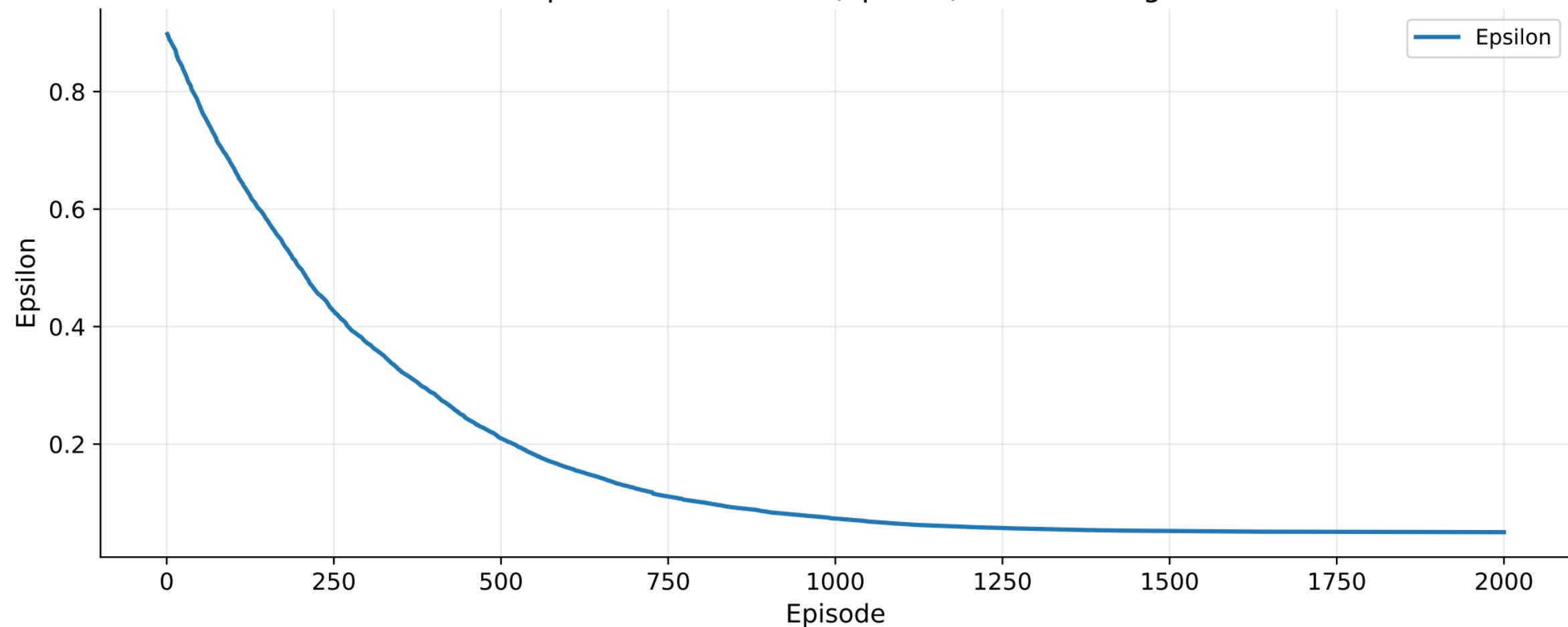
Final distance to goal over training



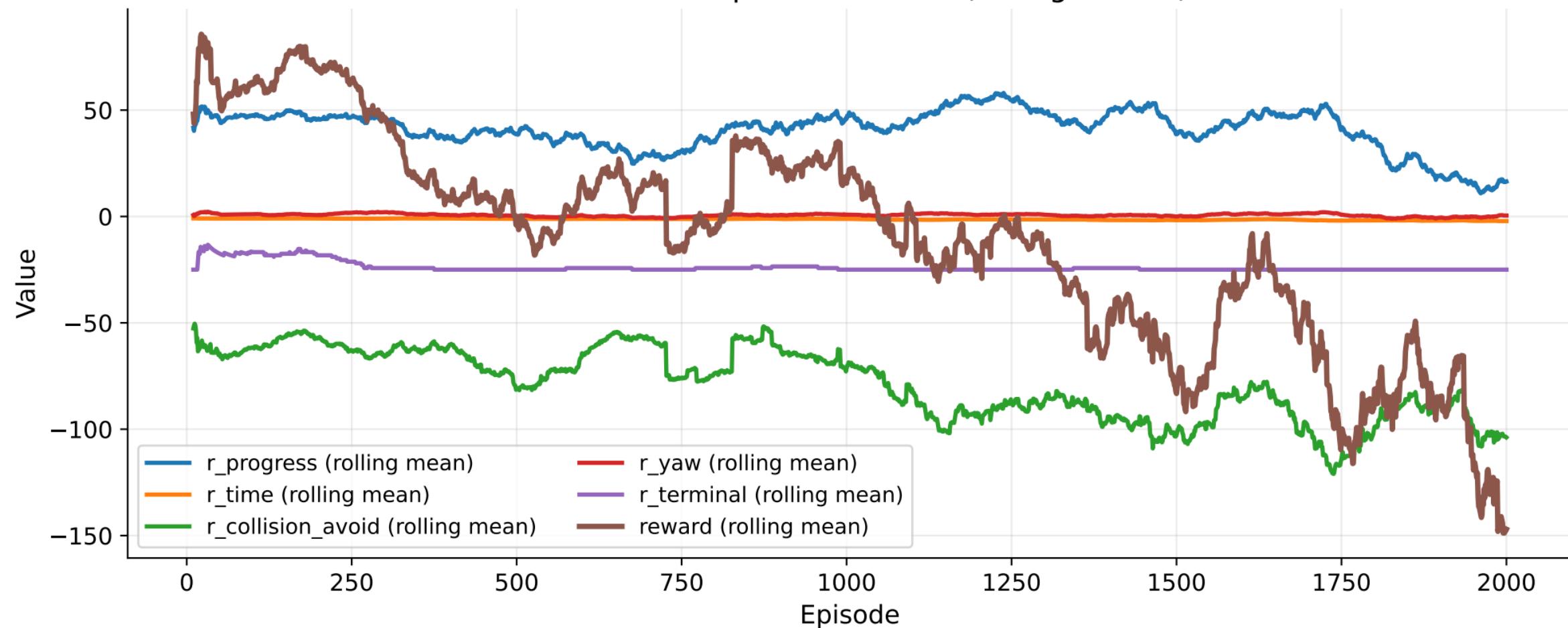
Minimum obstacle distance (proxy for safety) over training



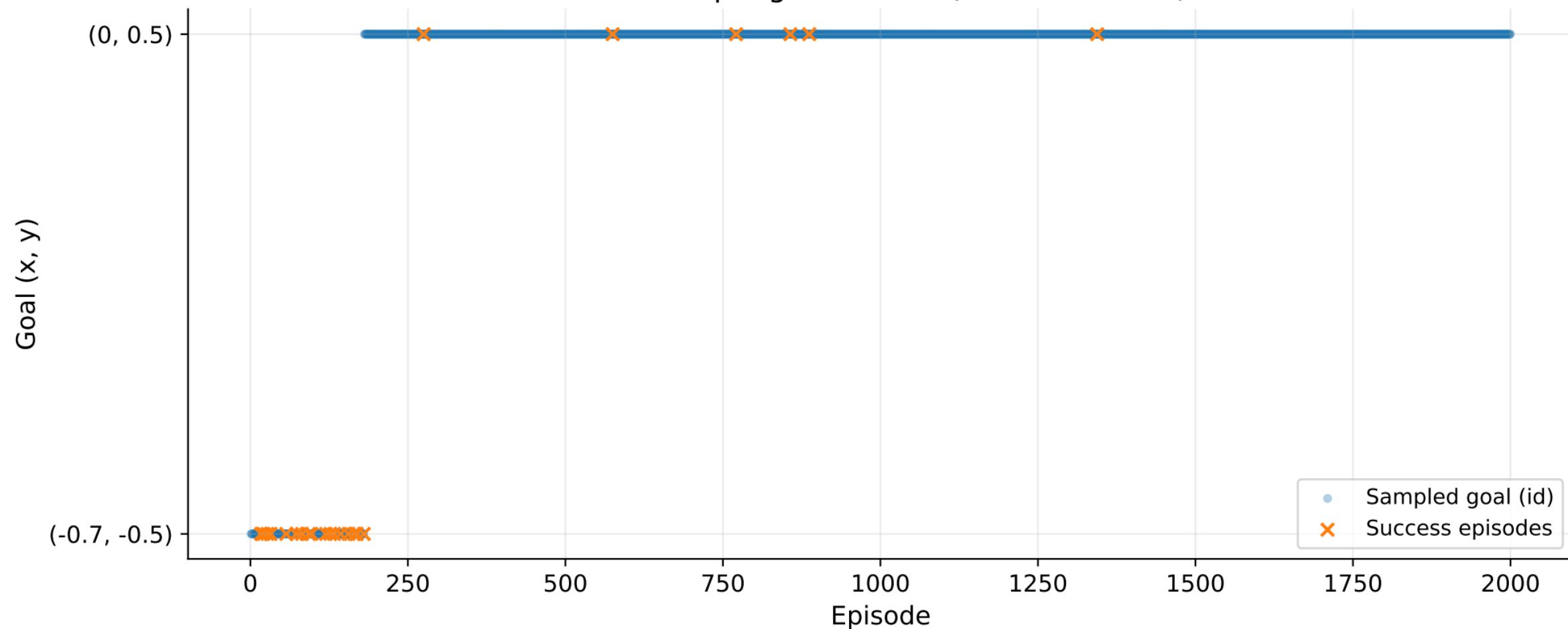
Exploration schedule (epsilon) over training



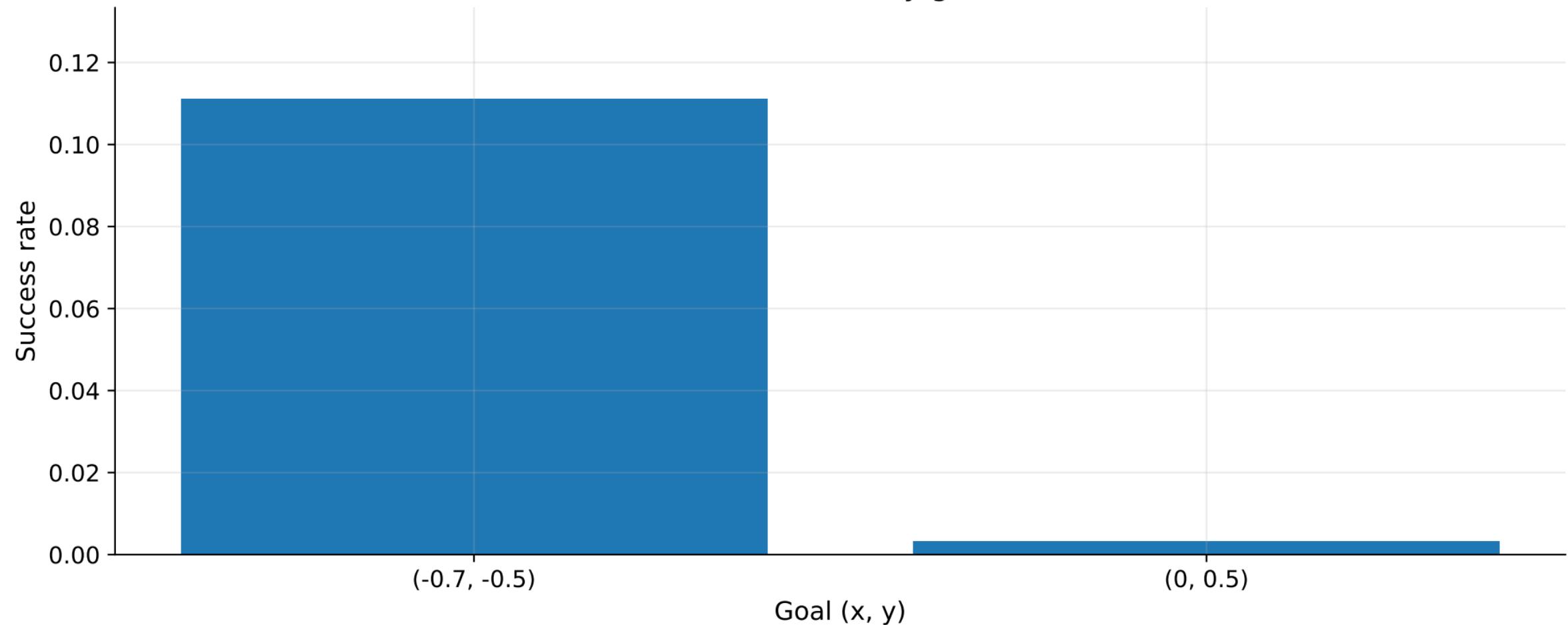
Reward decomposition trends (rolling means)



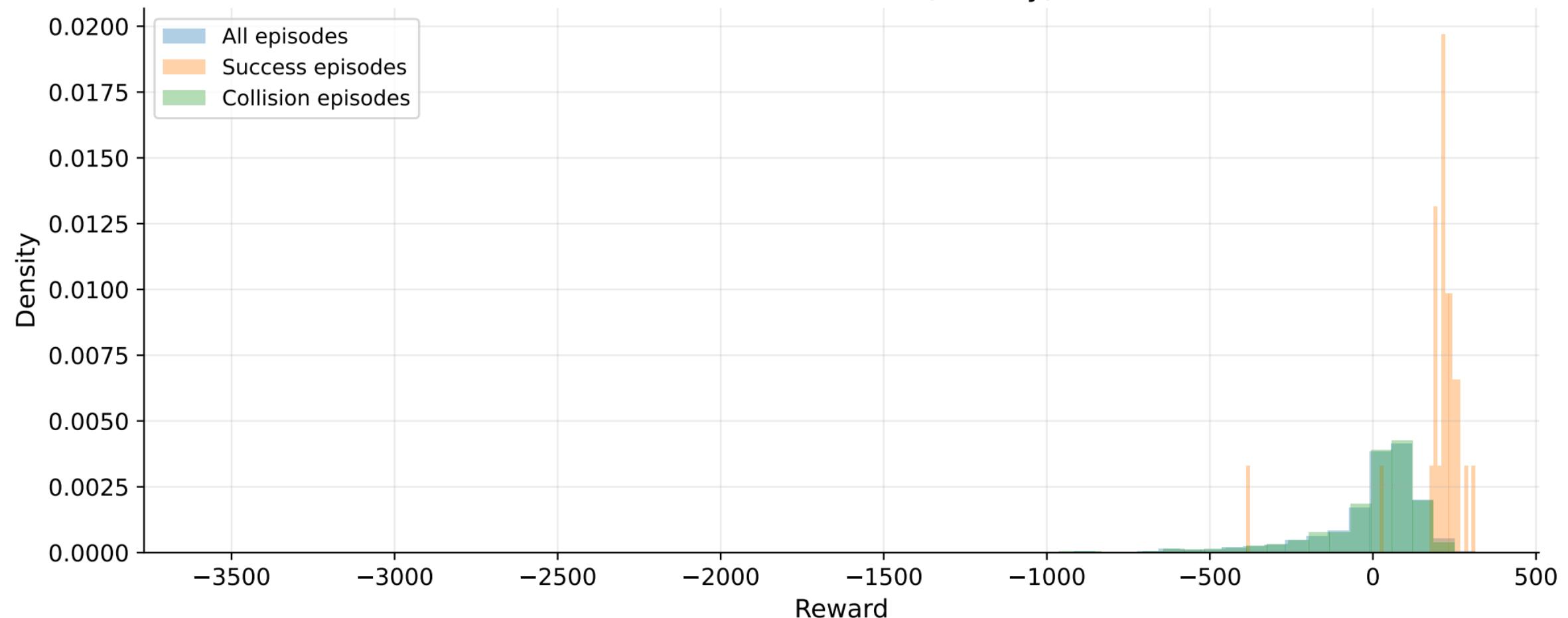
Goal sampling over time (and successes)



Success rate by goal



Reward distribution (density)



Consistency check: reward vs logged components

