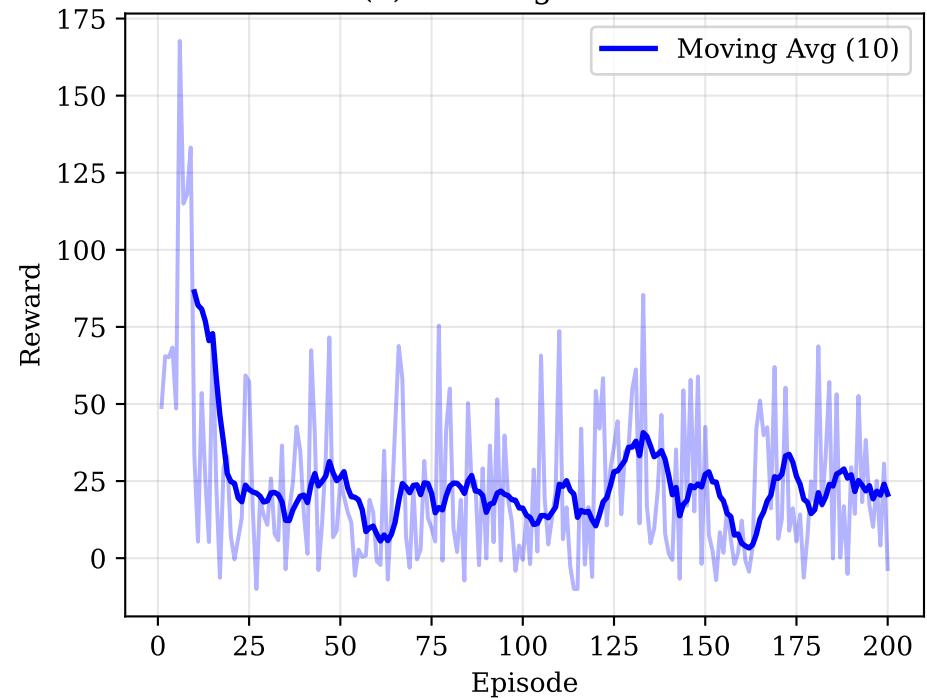
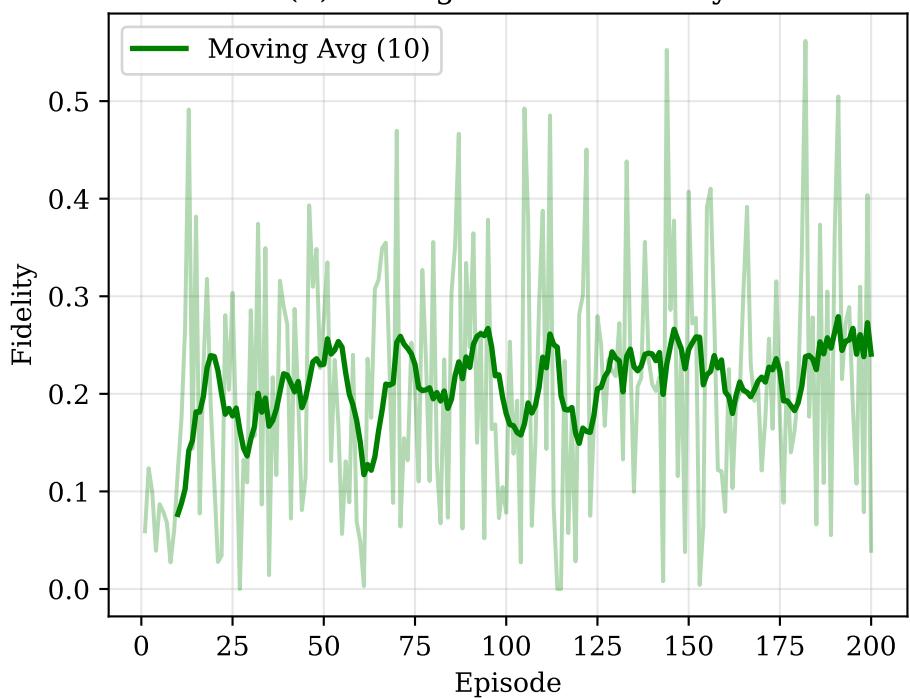


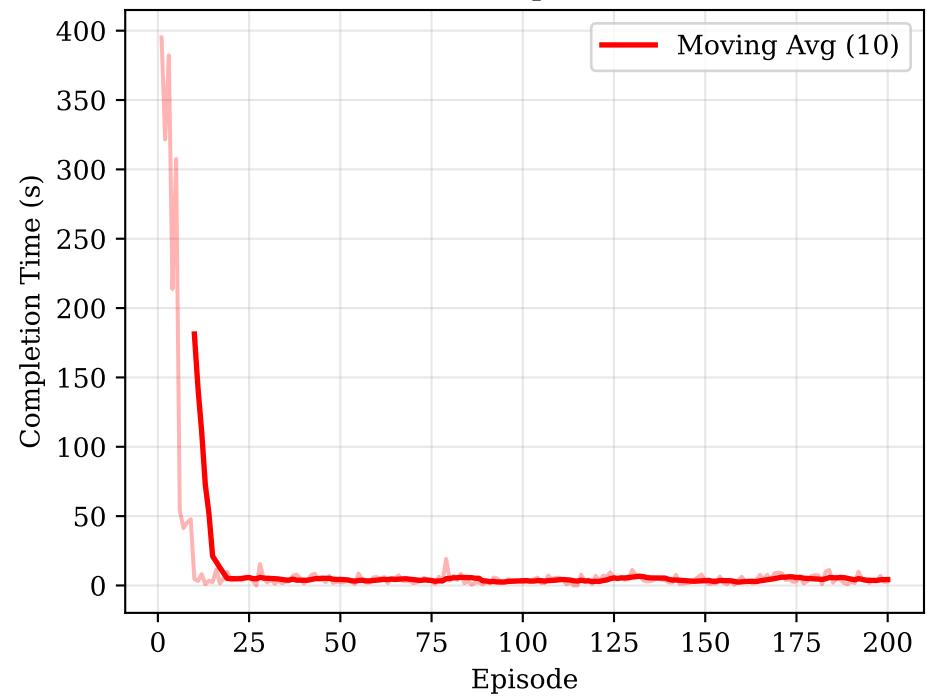
(a) Training Reward



(b) Average Circuit Fidelity



(c) Total Completion Time



(d) Time-Fidelity Trade-off

