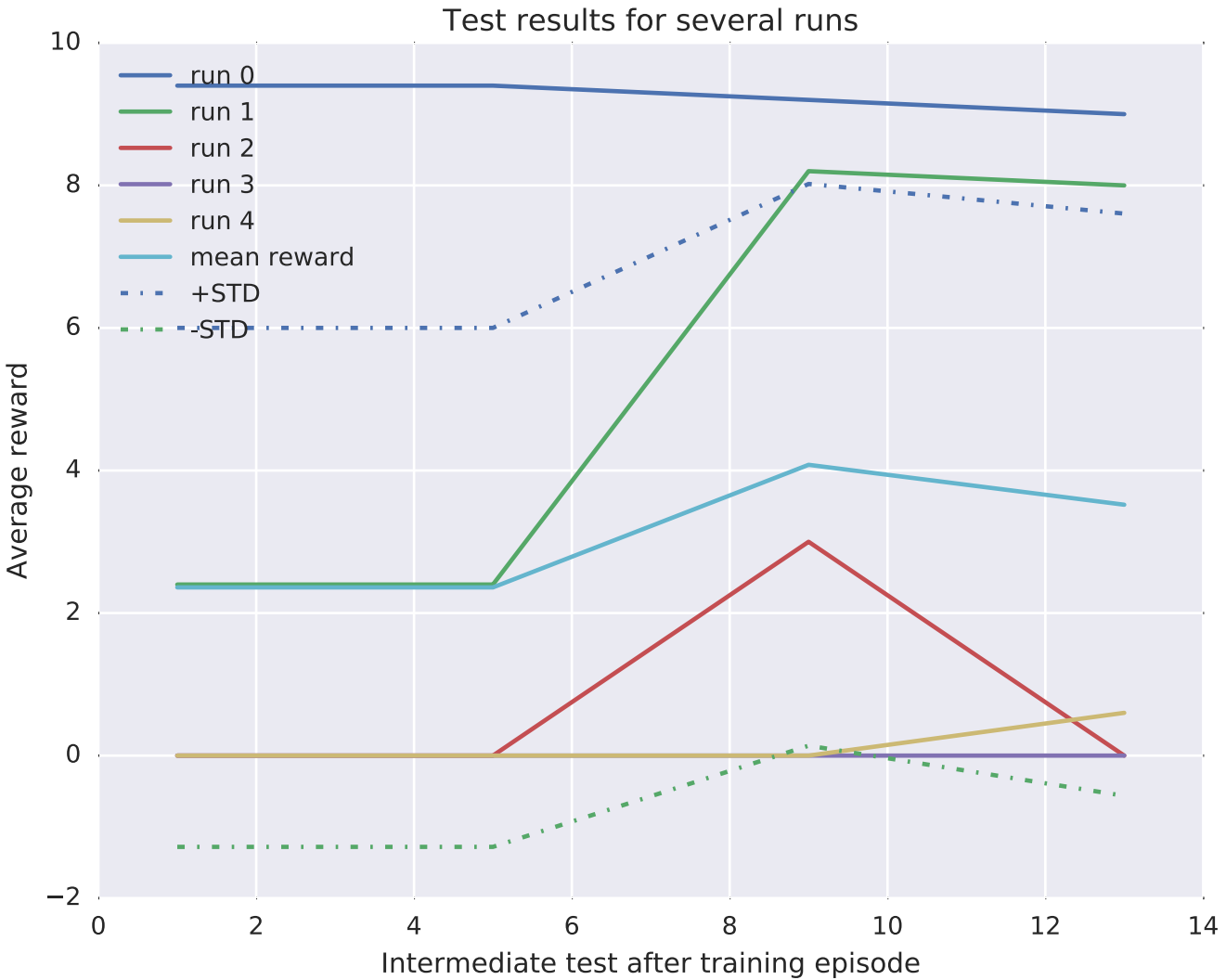
Test results for several runs
run 0
run 1
run 2
run 3
run 4
mean reward
+STD
-STD
Average reward
Intermediate test after training episode
10
8
6
4
2
0
−2
0
2
4
6
8
10
12
14