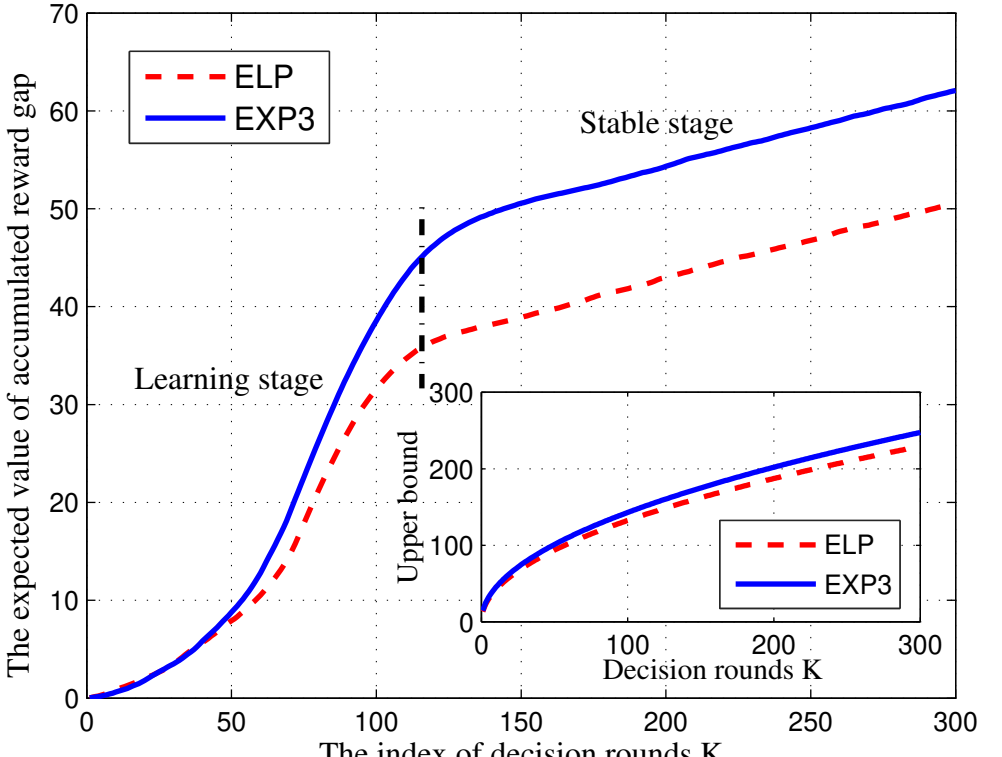
ELP
EXP3
Stable stage
Learning stage
The expected value of accumulated reward gap
The index of decision rounds K
Upper bound
Decision rounds K
0
10
20
30
40
50
60
70
0
50
100
150
200
250
300
100
200
300