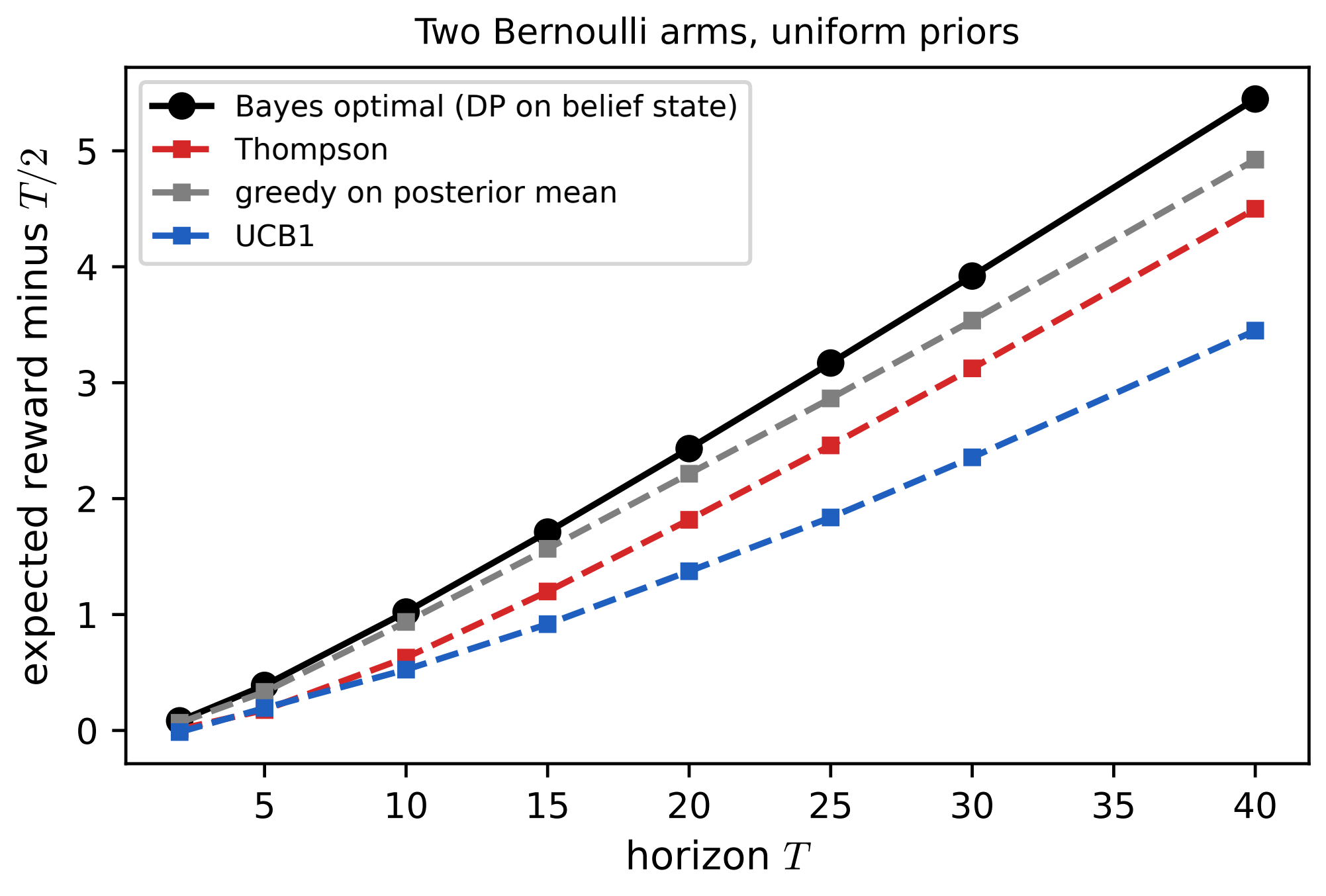

Two Bernoulli arms, uniform priors
Bayes optimal (DP on belief state)
Thompson
greedy on posterior mean
UCB1
expected reward minus $T/2$
horizon T
0
1
2
3
4
5
5
10
15
20
25
30
35
40