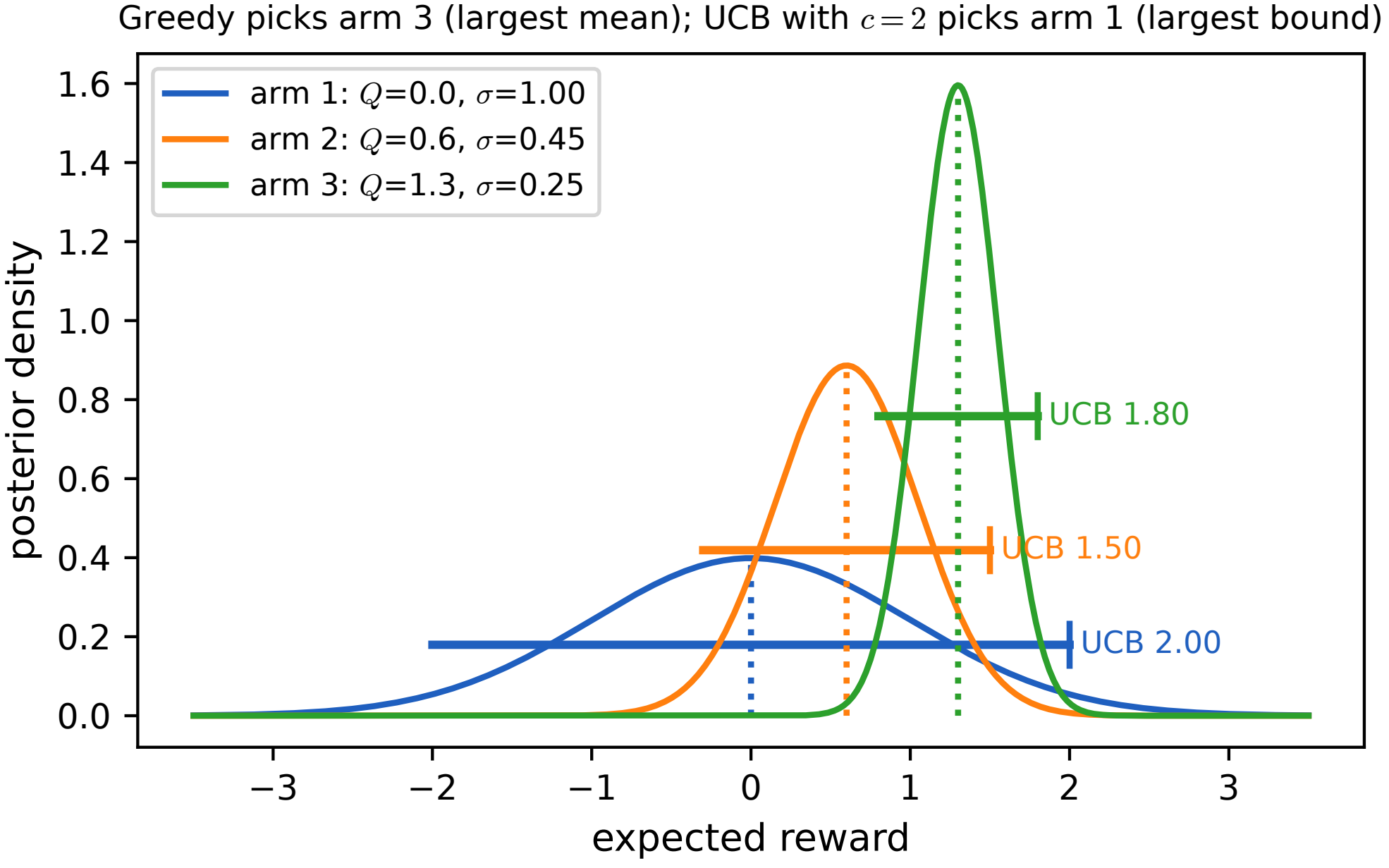

Greedy picks arm 3 (largest mean); UCB with $c=2$ picks arm 1 (largest bound)
arm 1: $Q$=0.0, $\sigma$=1.00
arm 2: $Q$=0.6, $\sigma$=0.45
arm 3: $Q$=1.3, $\sigma$=0.25
UCB 1.80
UCB 1.50
UCB 2.00
posterior density
expected reward
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6
−3
−2
−1
0
1
2
3