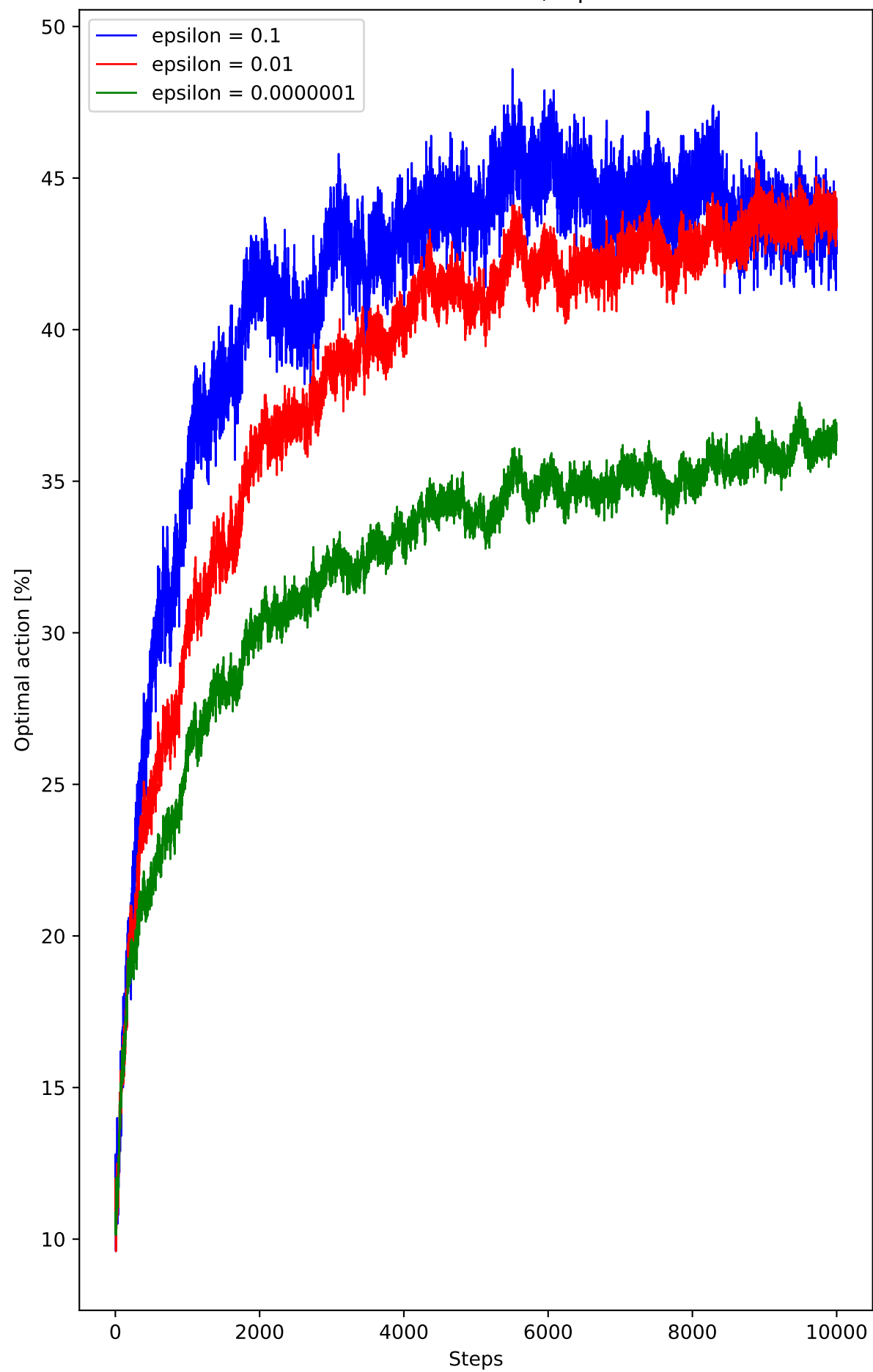


Optimal action for a 10-bandit algorithm  
with RW action values,  $\alpha = 1$



Optimal action for a 10-bandit algorithm  
with RW action values,  $\alpha = 0.1$

