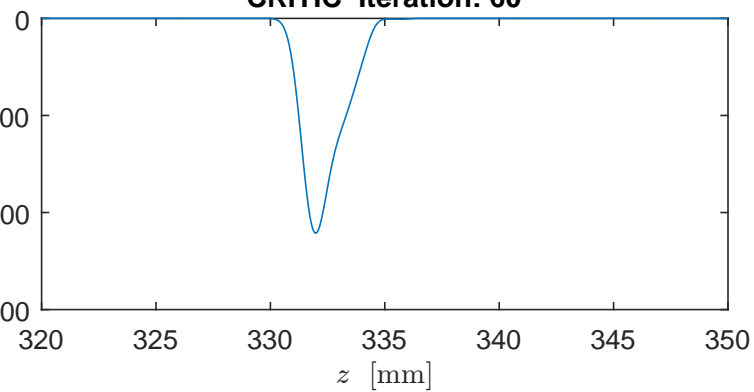
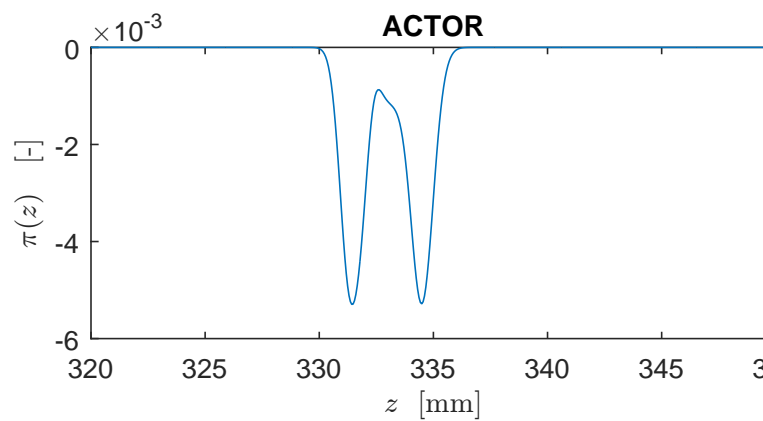


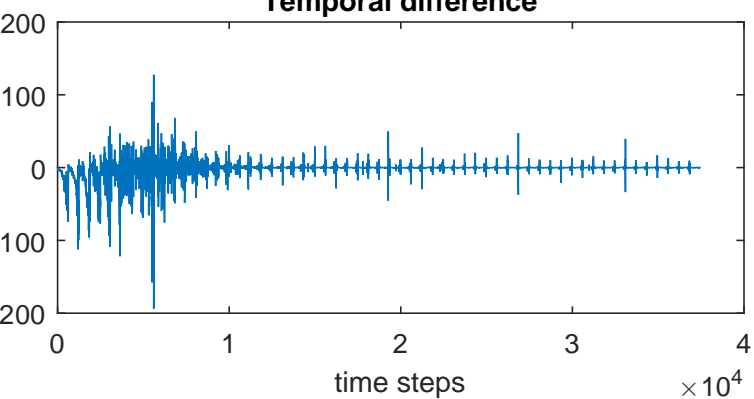
CRITIC Iteration: 60



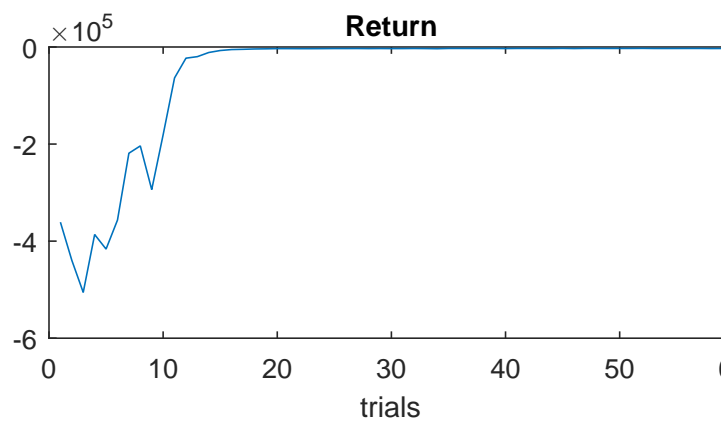
ACTOR



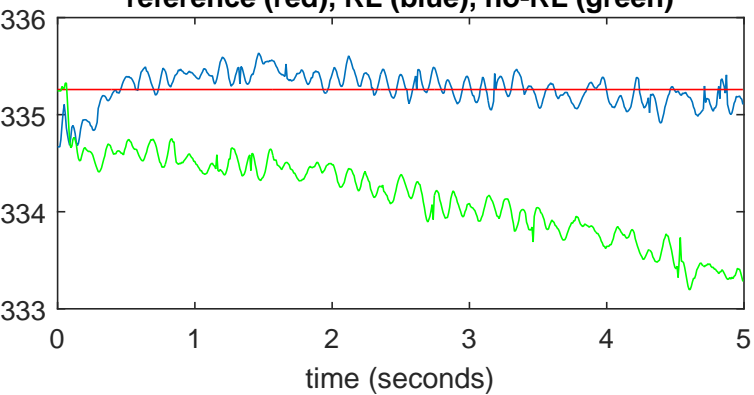
Temporal difference



Return



reference (red), RL (blue), no-RL (green)



Actor (b) & exploration (r)

