IDALAB

PARIS
LODRON
UNIVERSITÄT
SALZBURG

Simon Hirländer

Tutorial RL4AA

Incorporate considerations for safety

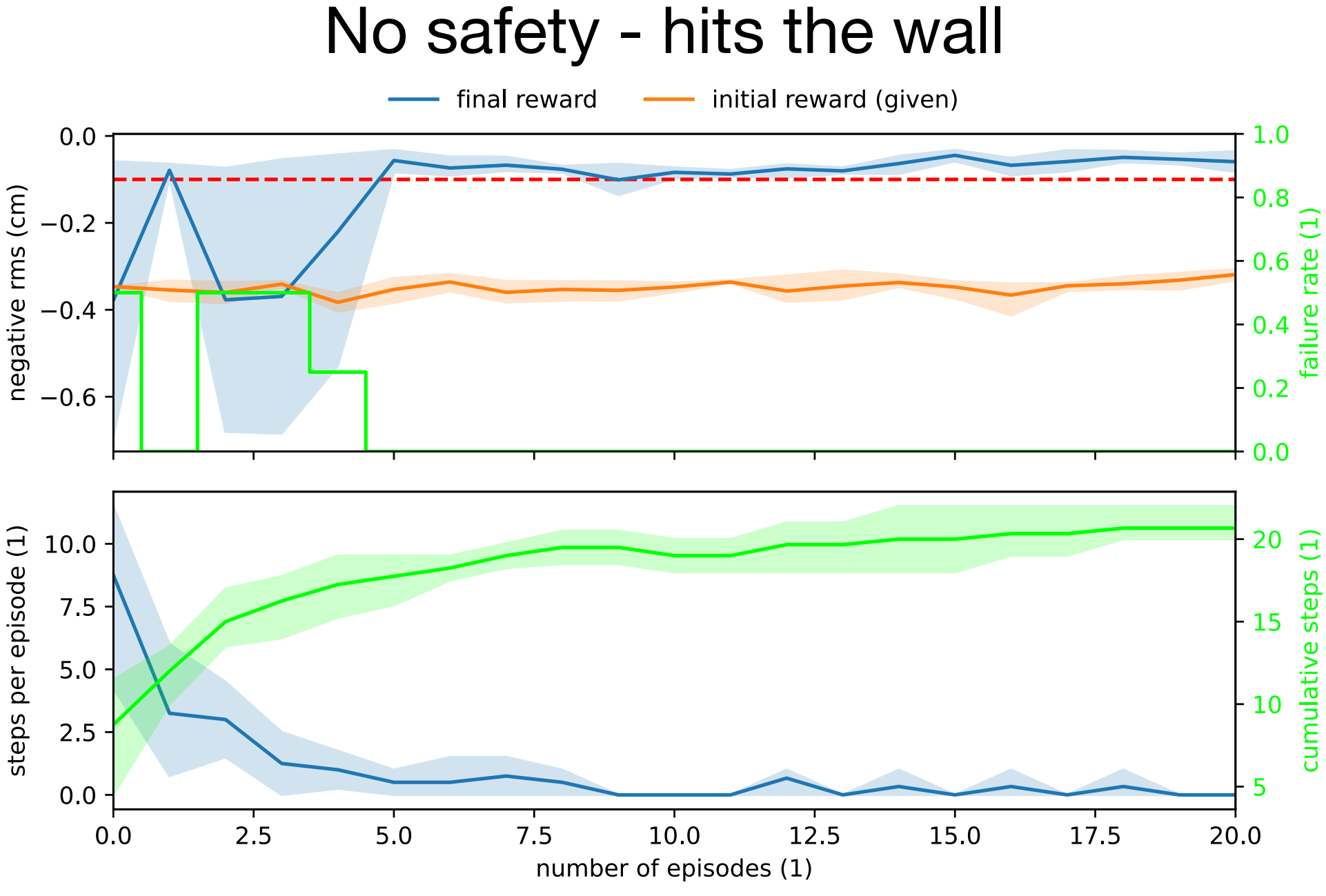
- Try to avoid hitting the wall

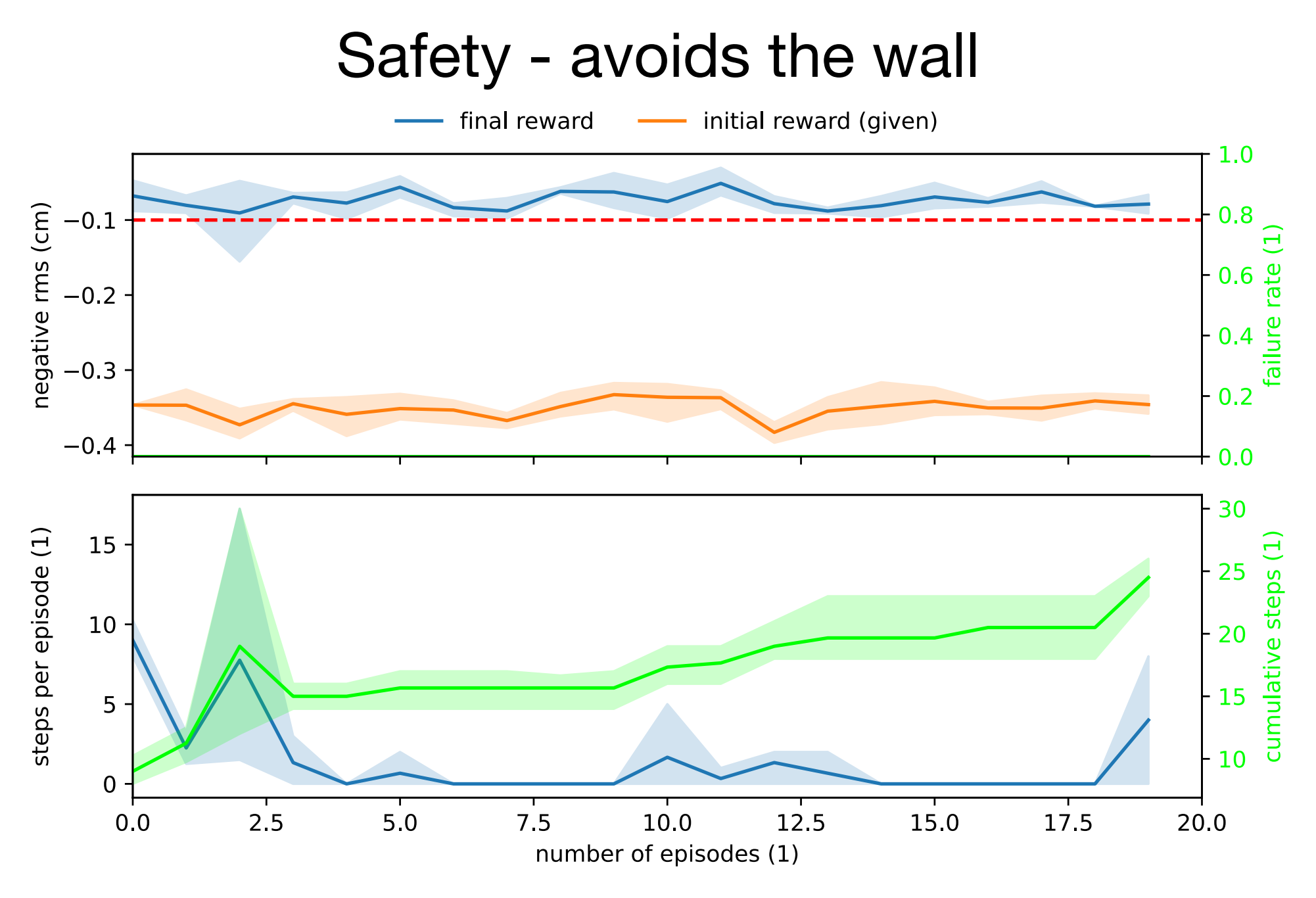
- Chance constrains:
  $\mathbb{P}(|s| > \text{threshold}) \geq \varepsilon \rightarrow$ safe policy is activated (red shaded)

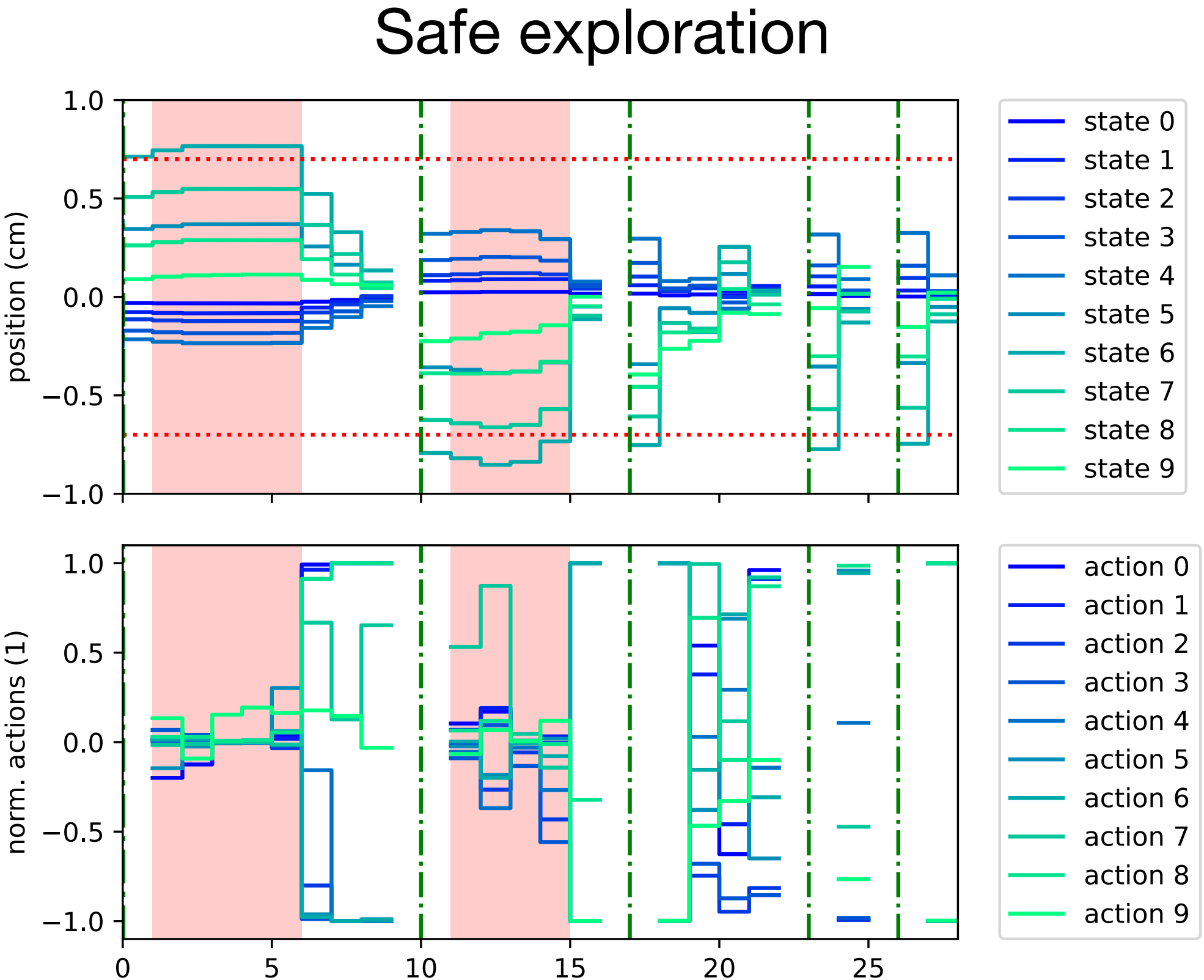
- Two layer safety: longterm safety (for optimal control) and instant safety (for safe exploration)

- Initial settings close to wall to test safeness

53

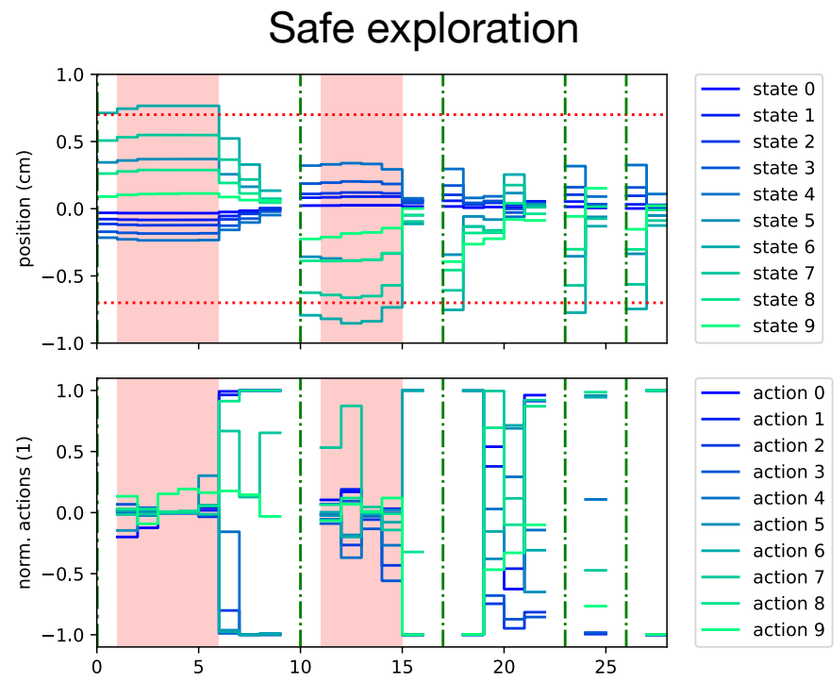

No safety - hits the wall
final reward
initial reward (given)
negative rms (cm)
failure rate (1)
steps per episode (1)
cumulative steps (1)
number of episodes (1)
0.0
−0.2
−0.4
−0.6
1.0
0.8
0.6
0.4
0.2
0.0
10.0
7.5
5.0
2.5
0.0
20
15
10
5
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0

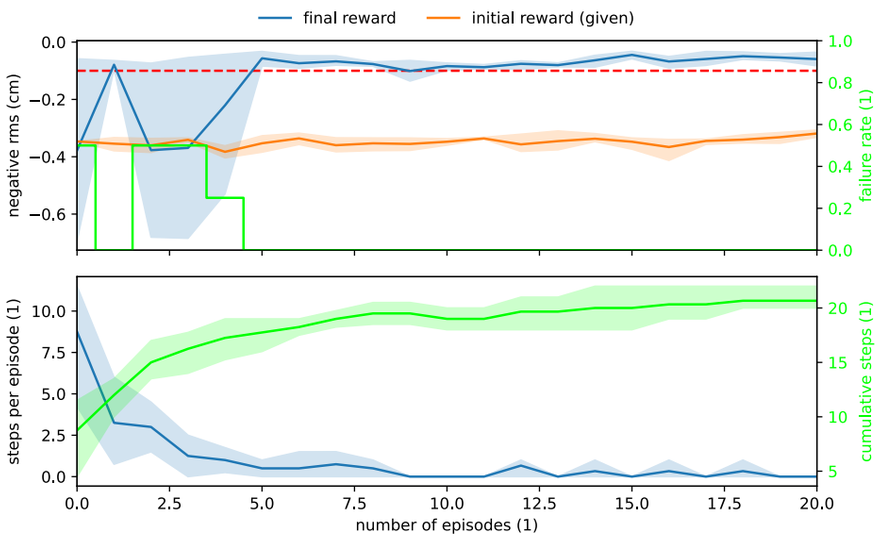

Safety - avoids the wall
final reward
initial reward (given)
negative rms (cm)
−0.1
−0.2
−0.3
−0.4
failure rate (1)
1.0
0.8
0.6
0.4
0.2
0.0
steps per episode (1)
15
10
5
0
cumulative steps (1)
30
25
20
15
10
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0
number of episodes (1)

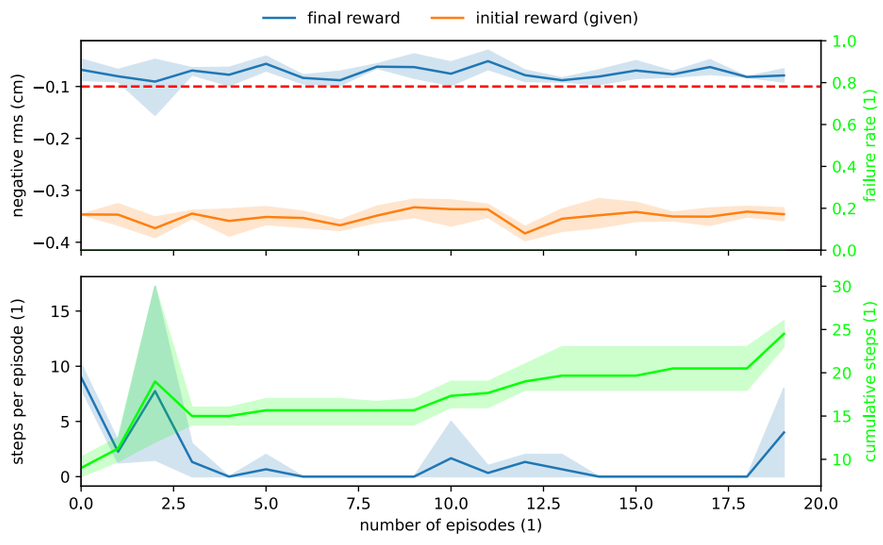

Safe exploration
position (cm)
norm. actions (1)
1.0
0.5
0.0
−0.5
−1.0
0
5
10
15
20
25
state 0
state 1
state 2
state 3
state 4
state 5
state 6
state 7
state 8
state 9
action 0
action 1
action 2
action 3
action 4
action 5
action 6
action 7
action 8
action 9

Incorporate considerations for safety

- Try to avoid hitting the wall
- Chance constrains:
  $\mathbb{P}(|s| > \text{threshold}) \geq \varepsilon \rightarrow$ safe policy is activated (red shaded)
- Two layer safety: longterm safety (for optimal control) and instant safety (for safe exploration)
- Initial settings close to wall to test safeness







Non stationarity and safety