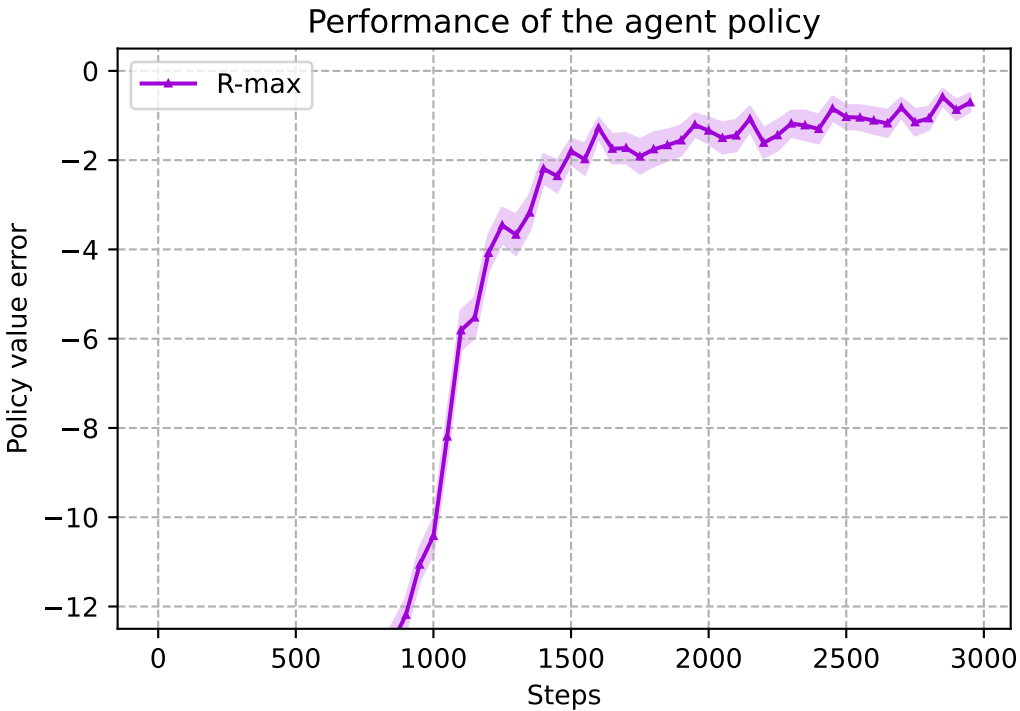
Performance of the agent policy
R-max
Policy value error
Steps
0
−2
−4
−6
−8
−10
−12
0
500
1000
1500
2000
2500
3000