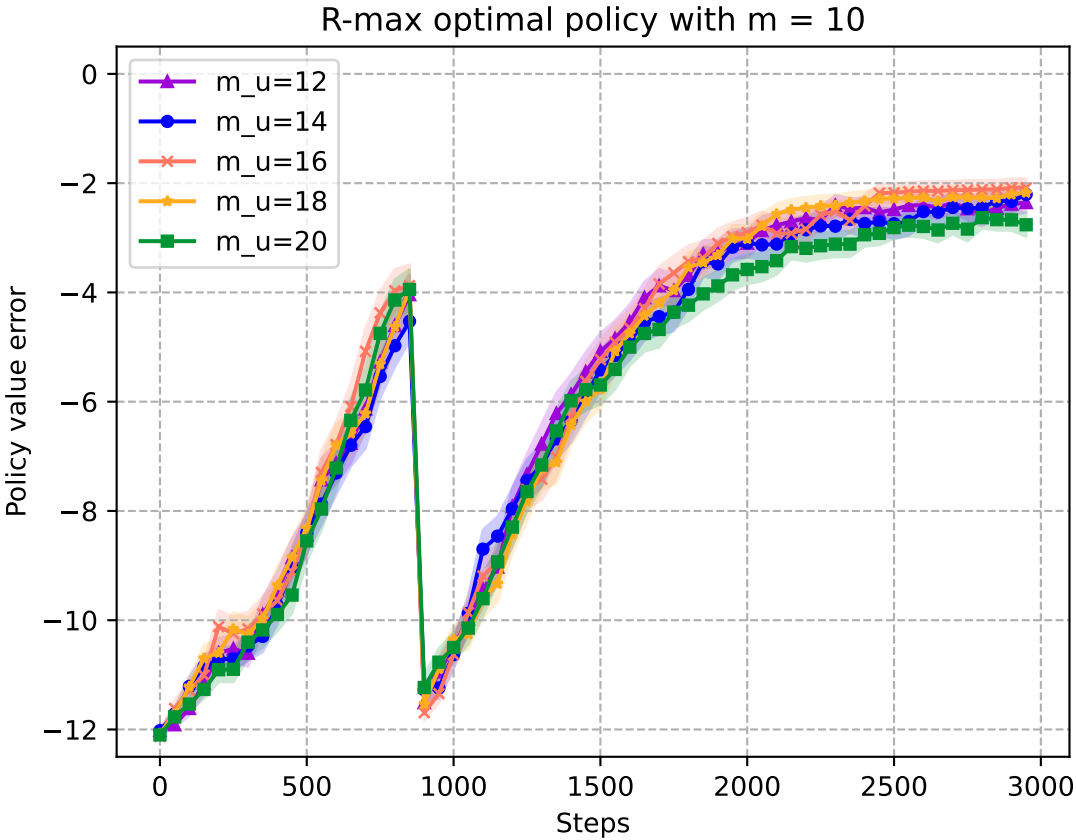
R-max optimal policy with m = 10
m_u=12
m_u=14
m_u=16
m_u=18
m_u=20
Policy value error
Steps
0
−2
−4
−6
−8
−10
−12
0
500
1000
1500
2000
2500
3000