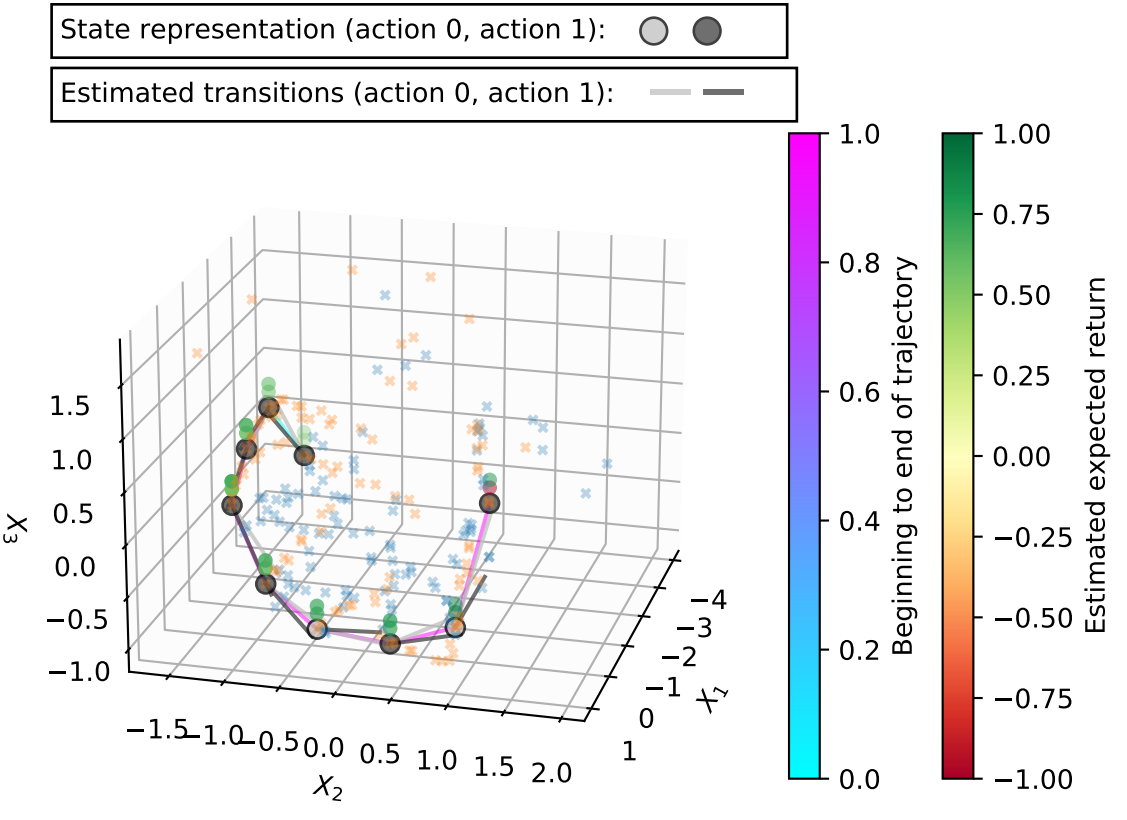
State representation (action 0, action 1):
Estimated transitions (action 0, action 1):
1.5
1.0
0.5
0.0
−0.5
−1.0
x_3
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
x_2
−4
−3
−2
−1
0
1
x_1
1.0
0.8
0.6
0.4
0.2
0.0
Beginning to end of trajectory
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
Estimated expected return