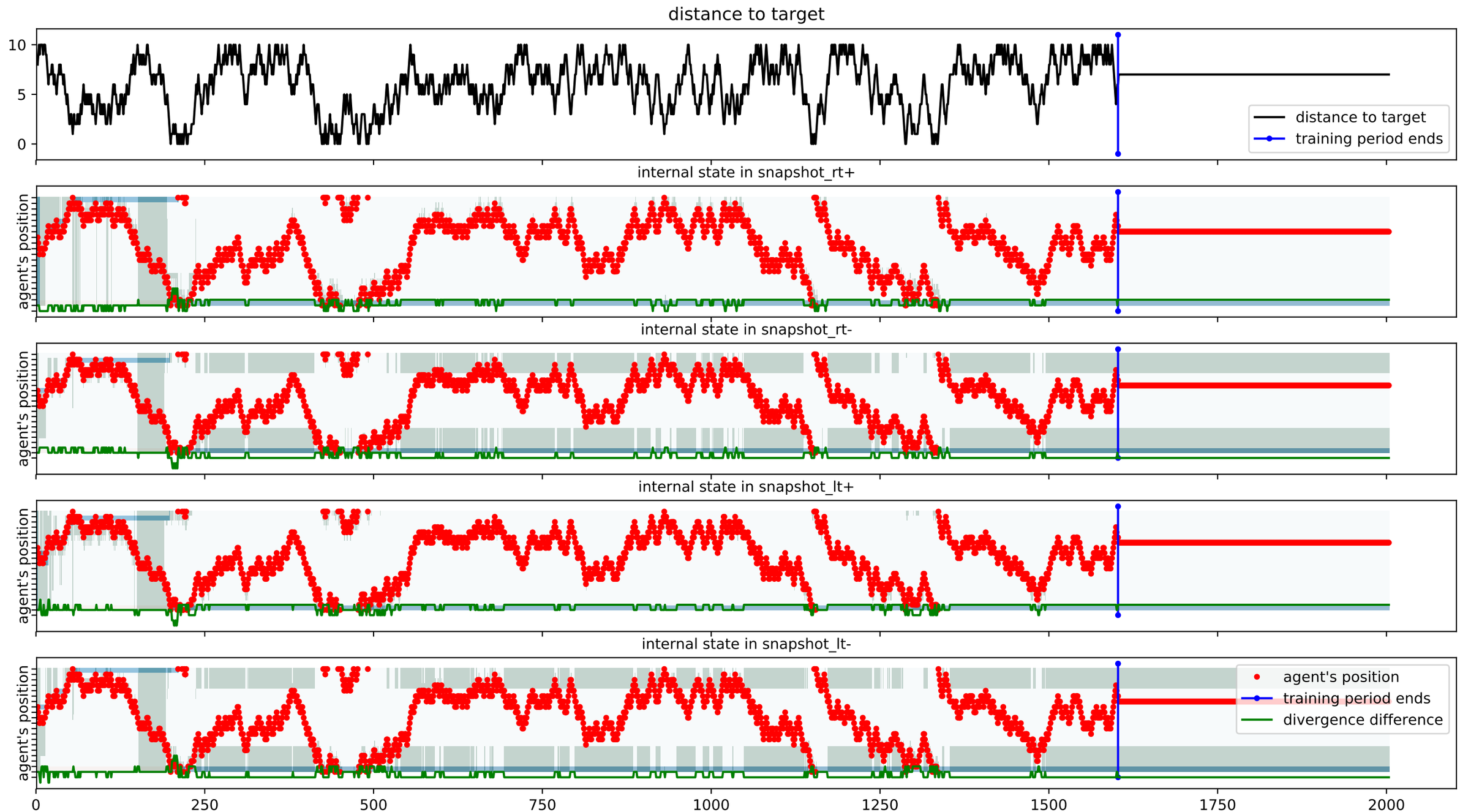
distance to target
10
5
0
distance to target
training period ends
internal state in snapshot_rt+
agent's position
internal state in snapshot_rt-
agent's position
internal state in snapshot_lt+
agent's position
internal state in snapshot_lt-
agent's position
agent's position
training period ends
divergence difference
0
250
500
750
1000
1250
1500
1750
2000