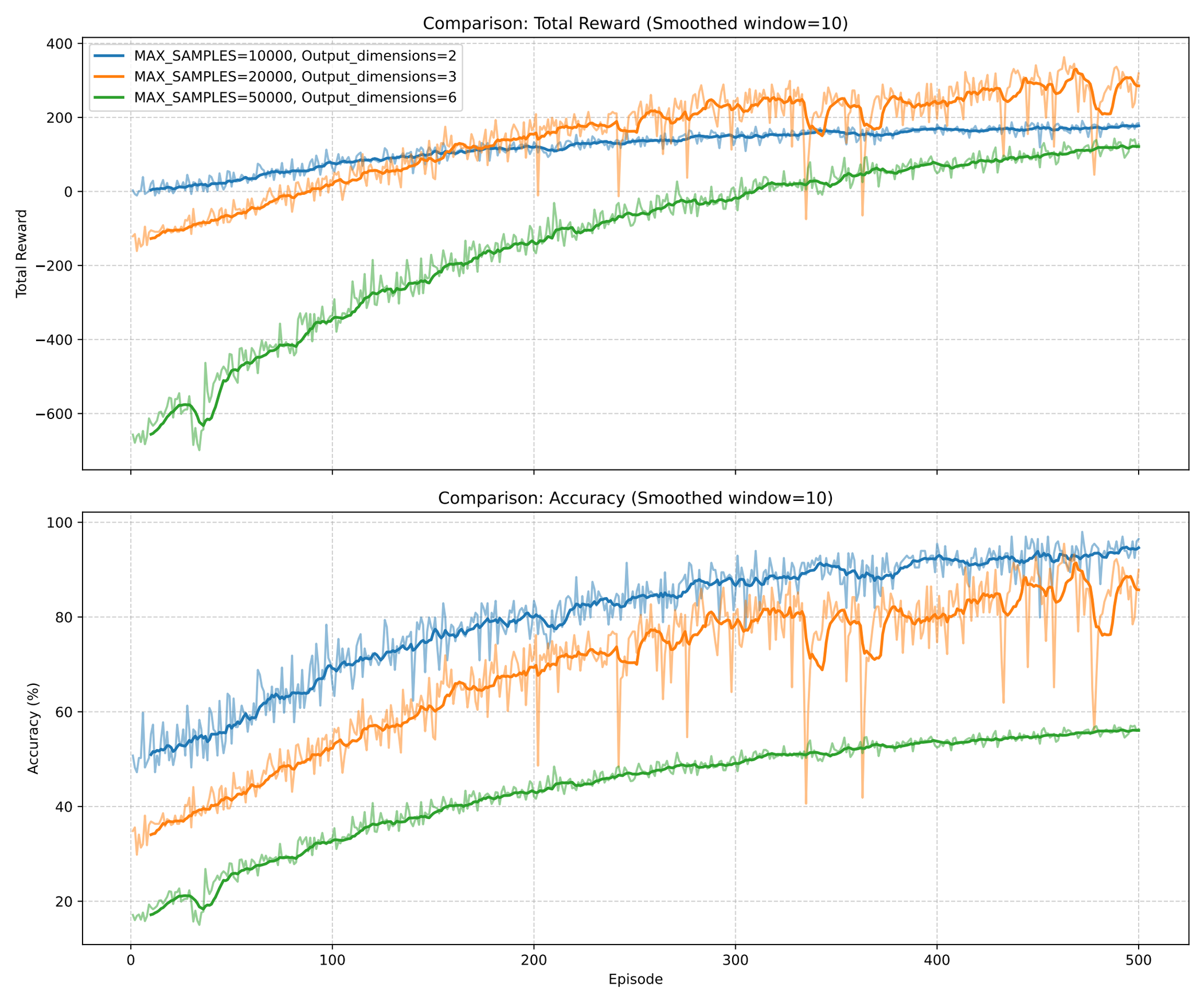
Comparison: Total Reward (Smoothed window=10)
MAX_SAMPLES=10000, Output_dimensions=2
MAX_SAMPLES=20000, Output_dimensions=3
MAX_SAMPLES=50000, Output_dimensions=6
Total Reward
400
200
0
−200
−400
−600
Comparison: Accuracy (Smoothed window=10)
Accuracy (%)
100
80
60
40
20
0
100
200
300
400
500
Episode