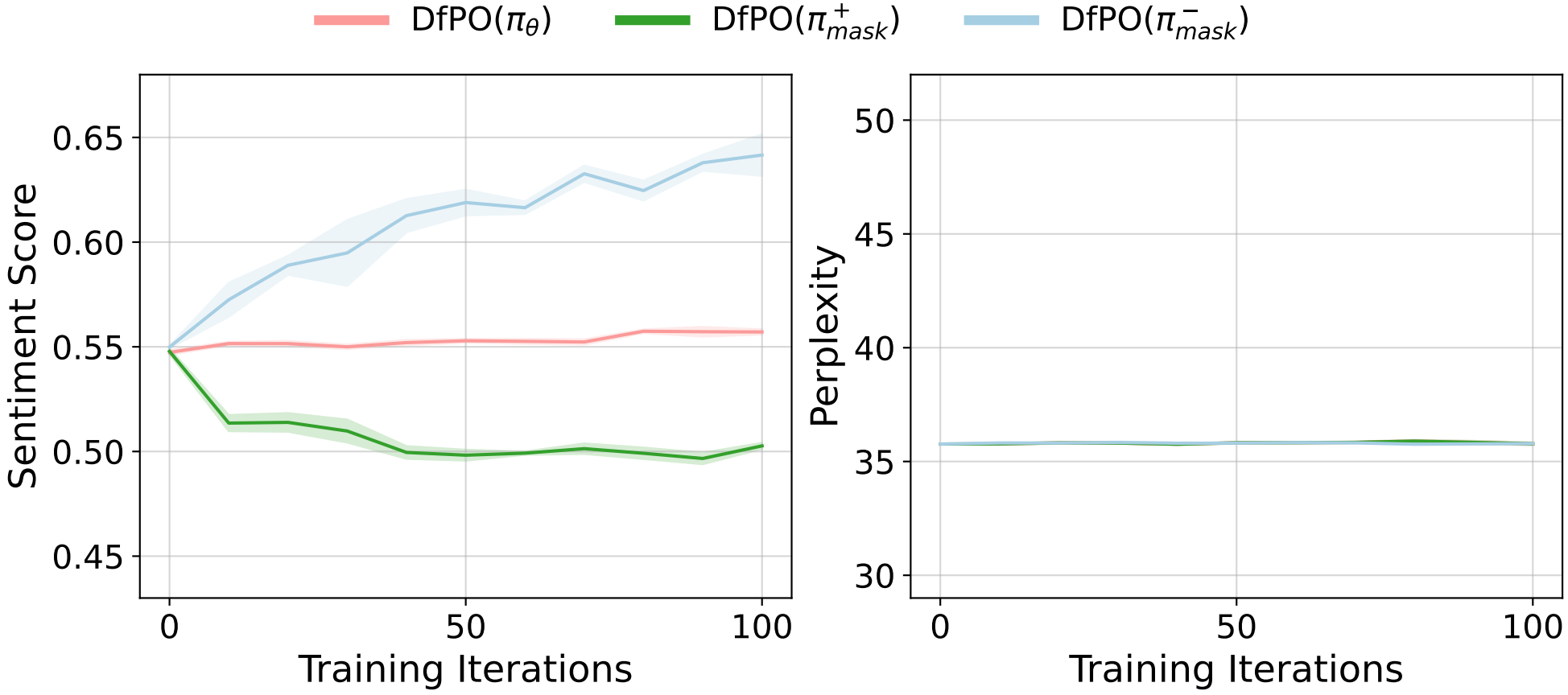

DfPO(π_θ)
DfPO(π^+_{mask})
DfPO(π^-_{mask})
Sentiment Score
0.65
0.60
0.55
0.50
0.45
0
50
100
Training Iterations
Perplexity
50
45
40
35
30
0
50
100
Training Iterations