Policy Gradient near π_2 0.30 **-** 7.29 7.28 0.25 (s_2, a_2) **-** 7.27 7.26 0.10 7.24 0.05

0.15

0.10

0.20

0.25

0.30

 π_2

0.05

7.23