

The figure is a line plot with 'Training episode' on the x-axis (0 to 200) and 'Training loss' on the y-axis (0.00 to 0.10). It displays the training loss for five individual runs (run 0 to run 4) and their statistical summary. Run 0 is blue, run 1 is green, run 2 is red, run 3 is purple, and run 4 is yellow. The 'mean of means' is shown as a light blue line, '+STD' as a blue dotted line, and '-STD' as a green dotted line. All runs show a similar trend: initial fluctuations between 0.01 and 0.04, followed by a sharp increase starting around episode 50, peaking near 0.09 around episode 80, and then settling into a high-frequency oscillation between 0.05 and 0.09 for the remainder of the training.

