Convergence (Model Average)

Legend: batch=10K, batch=1K, batch=100

Axes: loss (y-axis) vs seconds (x-axis)