Adam: the graph of absolute diff value varing with the number of iterations 2.0 train loss evaluation loss 1.8 1.6 1.4 1.2 1.0 0.8

150

Iteration

200

250

300

50

100