Epoch train log loss DY 0 train log loss CG-like Adam Ir=1e-3 Adam Ir=1e-3 CG-like Adam Ir=1e-4 Adam Ir=1e-4 CG-like Adam Ir=1e-5 -8 Adam Ir=1e-5 CG-like Adam Ir=1e-6 -10Adam Ir=1e-6 50 200 0 25 75 100 125 150 175 epoch