Epoch train log loss HS 0 train log loss CG-like Adam Ir=1e-3 Adam Ir=1e-3 -6 CG-like Adam Ir=1e-4 Adam Ir=1e-4 -8 CG-like Adam Ir=1e-5 Adam Ir=1e-5 CG-like Adam lr=1e-6 -10Adam Ir=1e-6 -1250 200 0 25 75 100 125 150 175 epoch