Epoch train log loss DY CG-like Adam Ir=1e-3 CoBA Ir=1e-3 CG-like Adam lr=1e-4 0 CoBA Ir=1e-4 CG-like Adam Ir=1e-5 **-**2 CoBA Ir=1e-5 train log loss CG-like Adam Ir=1e-6 CoBA Ir=1e-6 -6 -8 -1025 50 75 125 150 0 100 175 200 epoch