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

epoch