dim=256hidden_dim=128data_dim=64sample_size=128training_step=128r parallel initialization 0.6 gaussian initialization orthogonal initialization 0.5 0.4 0.3 0.2 0.1 0.0 -20 60 80 120 40 100

step