Evolution of loss sgd_train 1.75 sgd_test adahessian_train 1.50 adahessian_test 1.25 1.00 0.75 0.50 0.25 0.00 50 100 150 200 250 350 0 300 epoch