accuracy=0.245 Ir=0.001, betas=(0.9, 0.999), weight decay=0.0005, hessian power=1 train 2.75 test 2.50 2.25 2.00 SO 1.75 1.50 1.25 1.00 0.75 5.0 10.0 0.0 2.5 7.5 12.5 15.0 17.5 epoch

Evolution of loss during training: renormalized, adahessian