

Evolution of loss during training: renormalized, adahessian

accuracy=0.245

lr=0.001, betas=(0.9, 0.999), weight_decay=0.0005, hessian_power=1

