

MSE for plain SGD (varying  $\eta$ , 250 epochs, batch size: 50)

