

MSE for SGD (varying  $\eta$ , 100 epochs, batch size: 5)