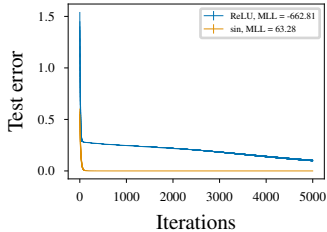
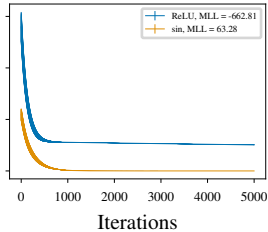


gradient descent w/Adam,  $\eta = 0.003$



gradient descent w/Adam,  $\eta = 0.0003$



vanilla gradient descent,  $\eta = 0.01$

