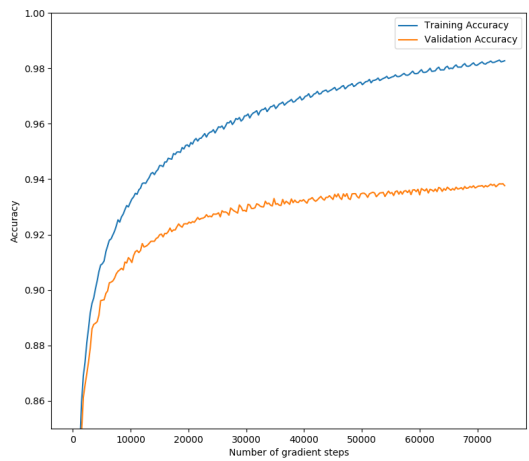
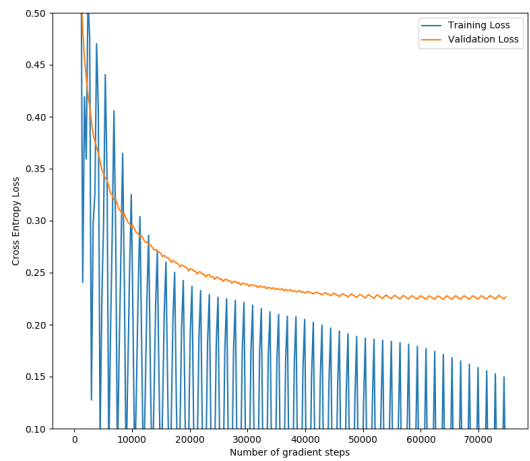


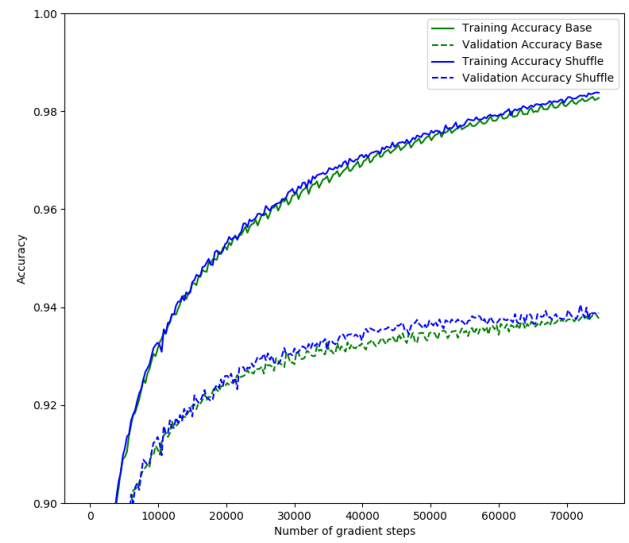
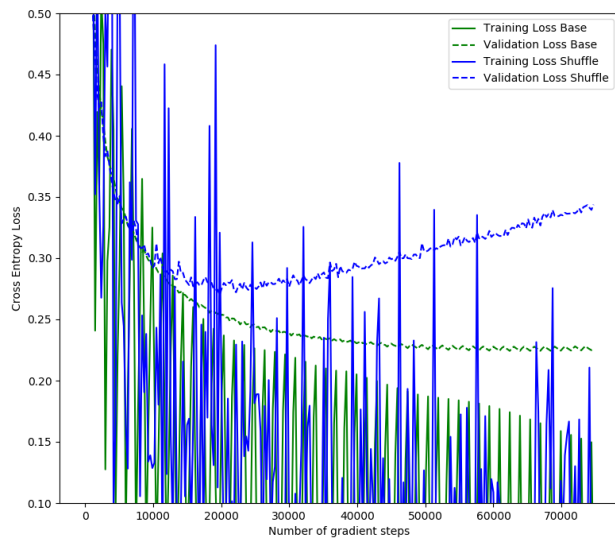
2c)



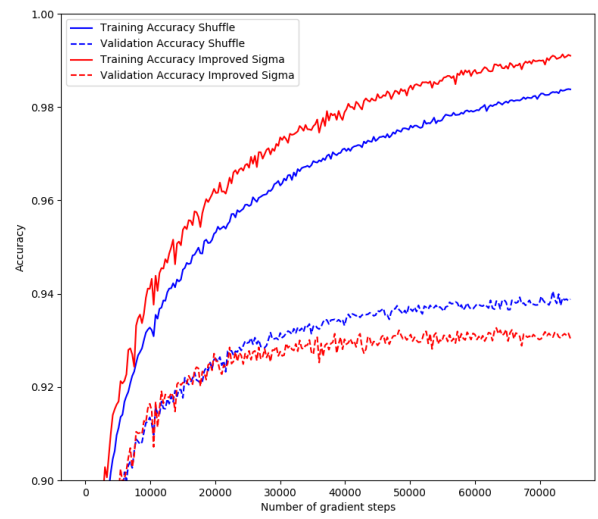
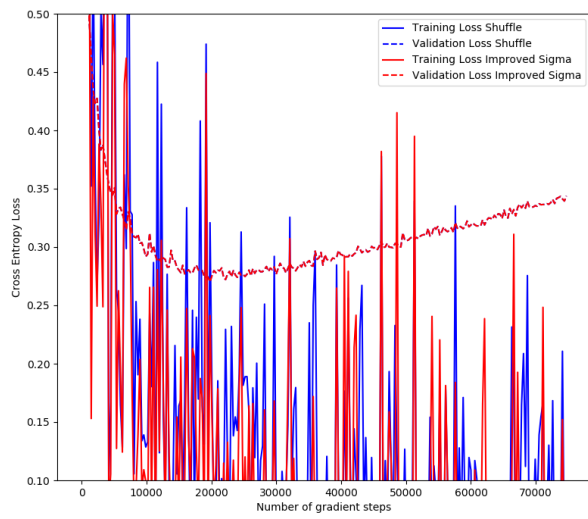
2d)

As we are using the bias trick all the weights and biases are included in the weight matrices. There are two weight matrices with the dimensions  $[786, 64]$  and  $[64, 10]$ . In total there are  $786 * 64 + 64 * 10 = 50944$  parameters.

3a)



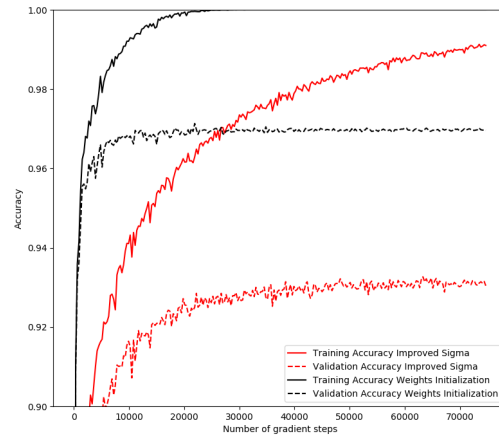
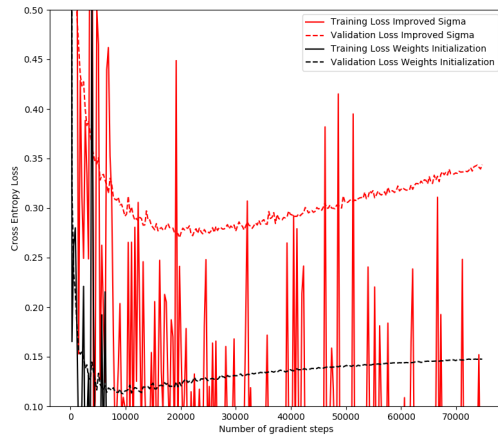
3b)



3c)

Fan-in:

- First Layer = 785
- Second Layer = 64



3d)

