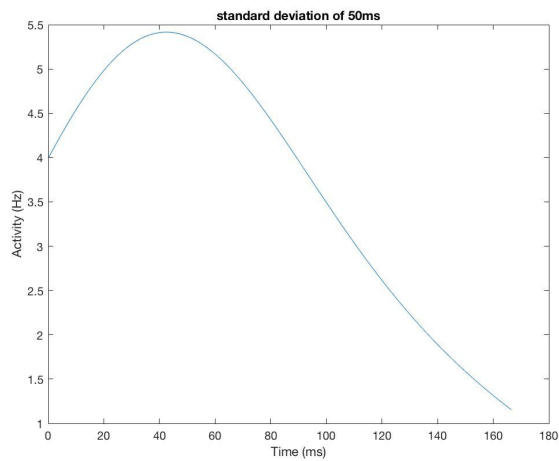
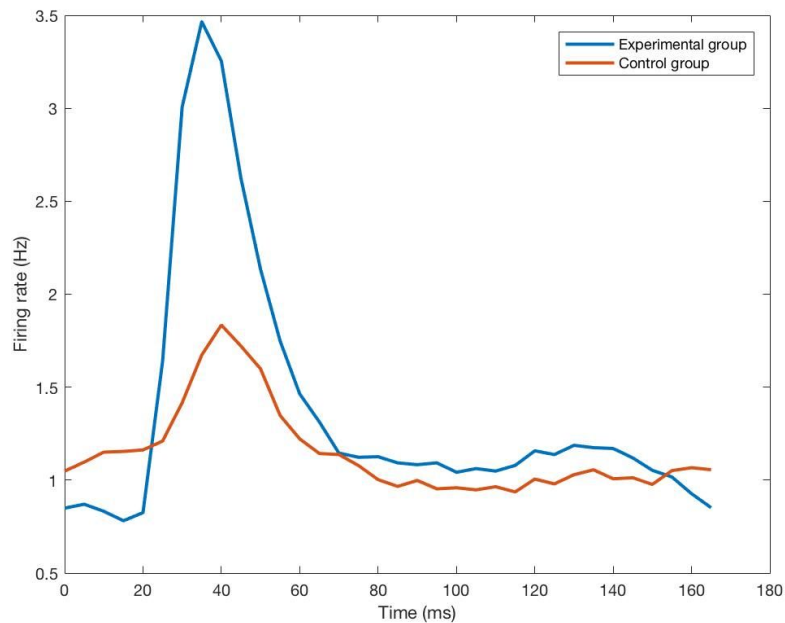


(c)

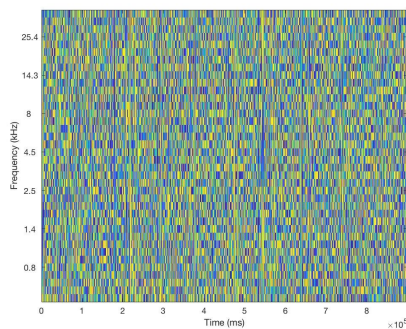


As shown above, a small standard deviation produces a curve with more specificity and variability, whereas a larger standard deviation creates a wider curve that barely resembles the first plot.

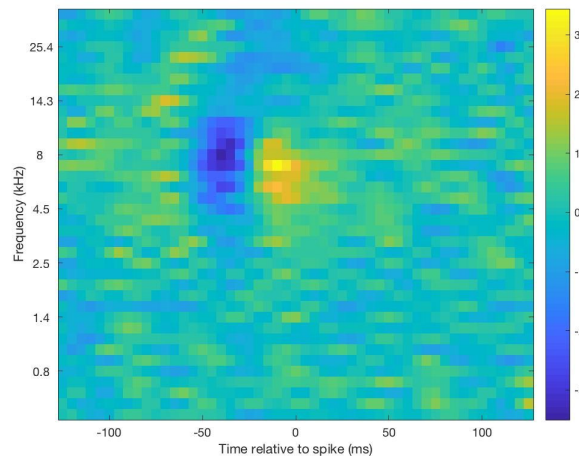


(d)

(e) There is a much greater spike peak in the firing rate from the experimental group at ~ 35 ms (about 2x the frequency), so it seems like neurons become more selective to the exposure stimulus. However, after the initial peak the differences in responses between groups are much smaller.

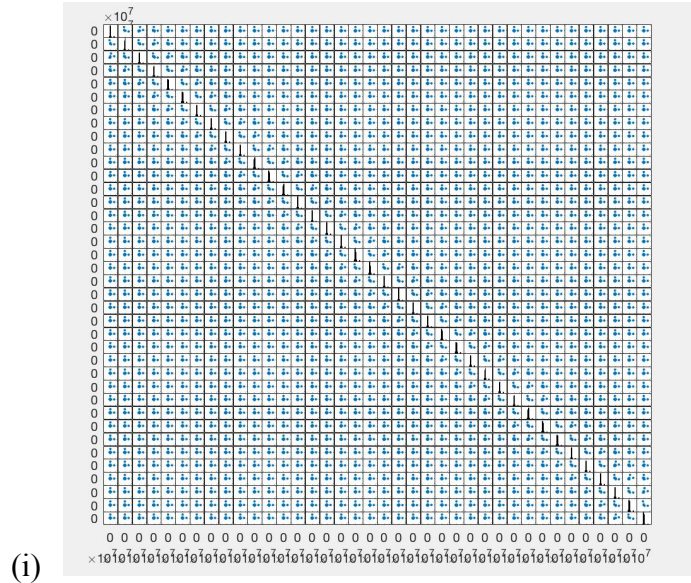


(f)



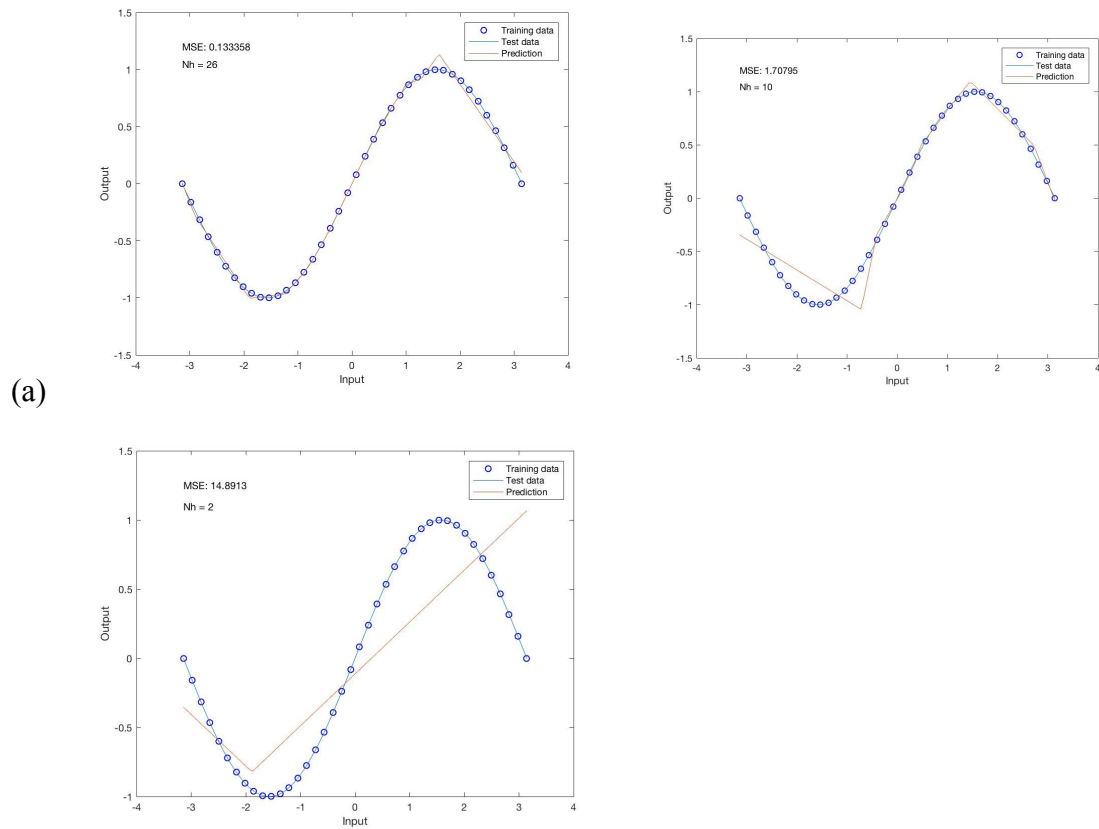
(g) ~ 7 kHz

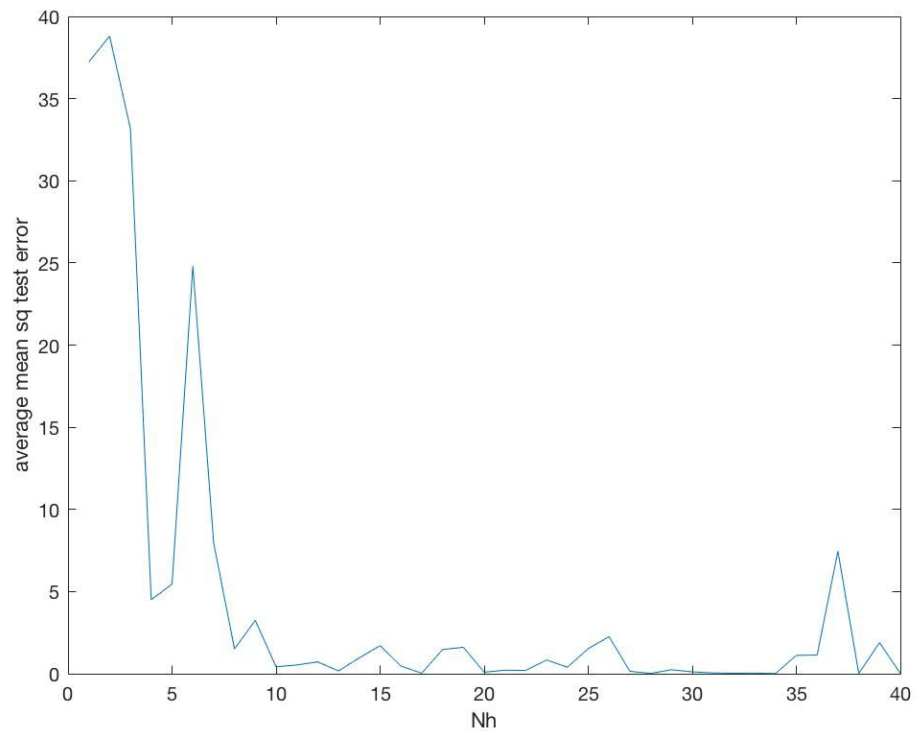
(h) The neuron's peak response would be higher to a brief tone pip, as shown by the plot. Based on the plot, the tone pip would be about 10 - 15 ms to elicit the largest response.



This does indeed resemble the identity matrix as you can see “ones” along the diagonal and extremely small values as the other entries in the matrix.

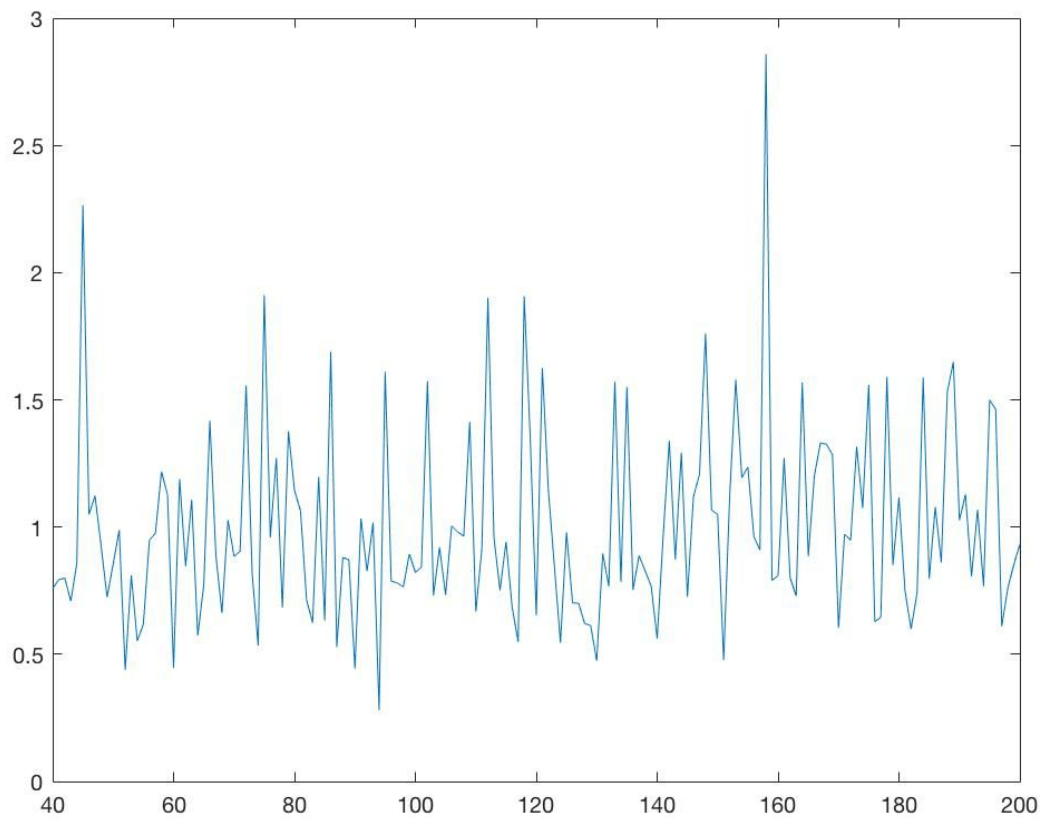
Problem 2





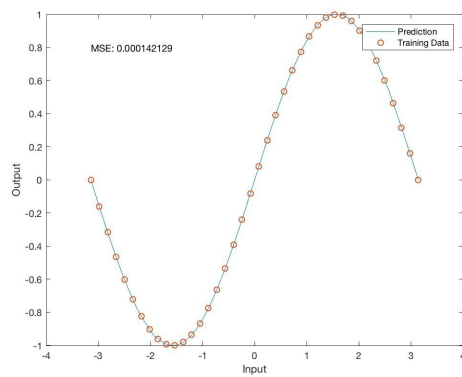
(b)

Best size model probably around $N_h = 20 - 22$. When the model is too small, the average MSE is extremely high. When it's too big, the average also gets somewhat high.



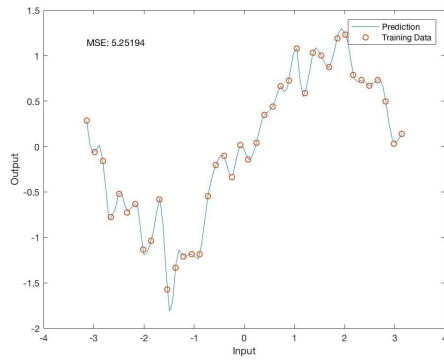
(c)

I tried the range from $nh = 40$ to 200 because it was taking too long with 500 as the upper limit. There seems to be a lot more variability in this bigger network, however the range is average MSE is much much smaller over a larger range of nh .



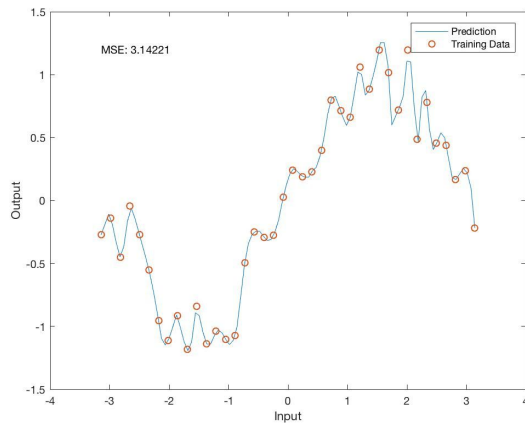
(d)

no label noise

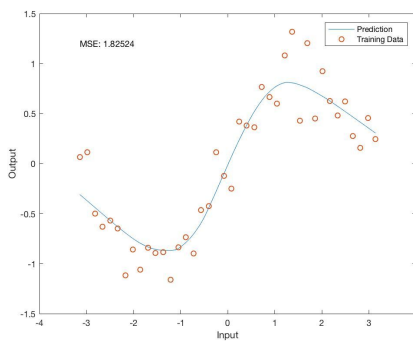


(e)

label noise 1



label noise 2



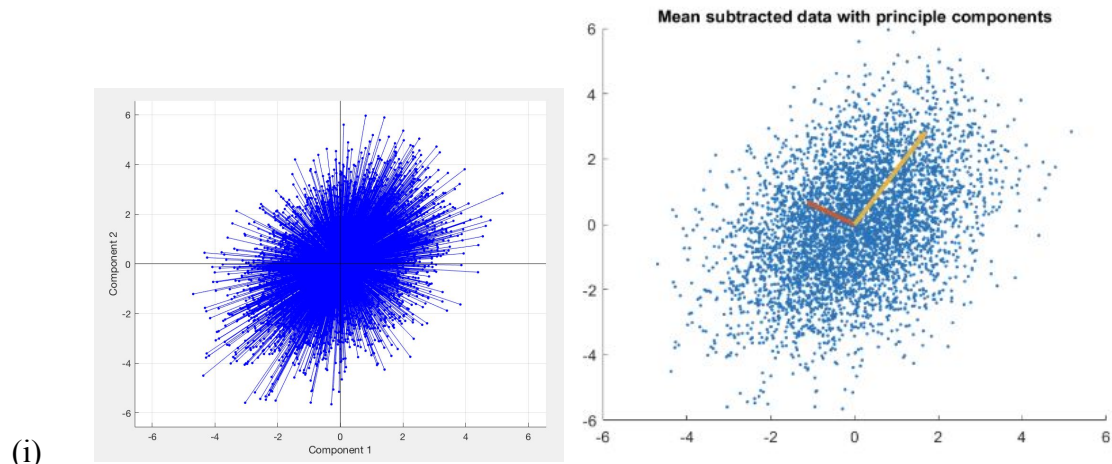
(f)

with lambda

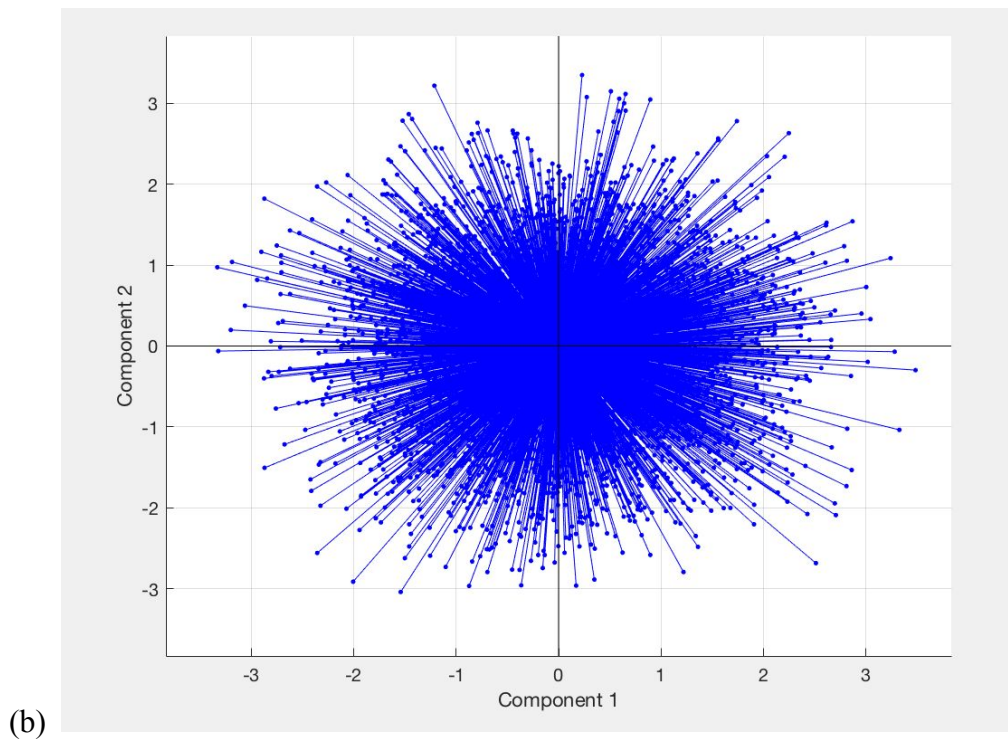
With label noise, MSE is between 3 and 6. Without label noise, MSE is extremely small (~ 0.0001). Thus, label noise appears to produce greater error. Adjusting the regularization parameter causes the MSE to remain between 1 and 2 (with or without label noise). So, I'm not sure if label noise is at all effective with an adjusted regularization parameter.

(g) Based on the results, it makes sense for neural networks to expand the dimension of their inputs as error seems to narrow down and lessen. Label noise can be regulated. More neurons than training data seems to be effective.

Problem 3



(ii) Sparseness will greatly expedite processing of larger data.



As you can see the data has been whitened and reoriented along new principal axes. You can simply find the principal components by looking at the axes, in fact the axes are indeed the principal components.

(c)

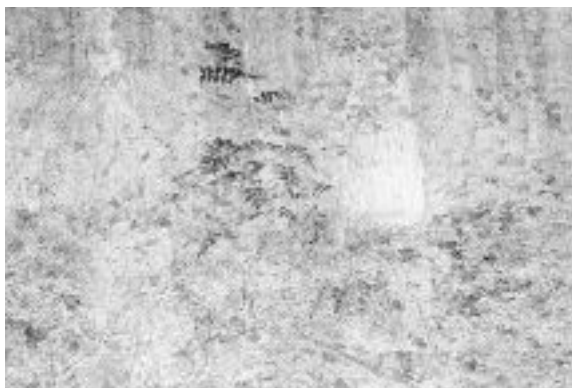


(i)





(ii)





(iii)





Wow! This actually worked! Magic. The images emerge with great clarity and distinctness. Learning following whitening makes everything much smoother than the other way around.