1 e) Clearly, the experimental group's neurons were more selective for the stimulus, firing both earlier and more rapidly when the stimulus tone was presented (though the control still responded). Interestingly, the experiment group's neurons fired more in general even during the brief amount of silence in the audio, indicating that either more neurons were being measured (e.g. the total number of neurons responsive to stimulus was the same and they just became more responsive) or more neurons became responsive to the stimulus.

g) The neuron appears to be most selective to frequencies around 8 khz

h) A tone pip, so as to avoid the inhibitory region before hand (similar to how just activating the center of a visual receptive field, avoiding the inhibitory regions, is needed to generate simple cell responses) would lead to the best response. Ideally the pip would be ~20ms long, the width of the exciting region

i) The spectrogram is fairly uncorrelated, though not completely. Around the main diagonal, the values are not very close to zero, indicating that pairs of frequencies close to each other are slightly correlated or anticorrelated

(1)

b) The best size is probably around 12 or 13, this seems to be one of the only stable spots in the graph, lots are remarkably spikey and changes every time I run the code even though I average over 100 trials

c) The mean squared error quickly drops off once the neural network size reaches around ~50 neurons, despite the regularization parameter being 0. The really big neural networks also have very low reliability with little variation between sizes and no large spikes, unlike with the smaller models that still suddenly explode even with averaging over large numbers of them.

(2)

b) The best size is probably around 12 or 13, this seems to be one of the only stable spots in the graph that is remarkably spiky and changes every time I run the code even though it averages over 500 trials

c) The mean squared error quickly drops off once the neural network size reaches around ~50 neurons, despite the regularization parameter being 0. The really big neural networks also have very low variability with little variation between sizes and no large spikes, unlike with the smaller models that still suddenly explode even with averaging over large numbers of them.

d) With noise and $\lambda = 0$, the large neural net badly overfits the data and has an MSE around 5. Once $\lambda = 0.0001$, however, it fits a much smoother curve with an MSE around 1-2.

e) It makes sense for neural nets to massively expand the dimension of their inputs because, not knowing the type of data they'll receive, it would be hard to consistently hit the sweet spot of network size, but using a massive network with some form of regularization would allow the brain to both compensate for noise on small datasets and learn well on a wide range of task complexities/dimensionalities. This is a scenario that the brain finds itself in fairly often, I'd assume, such as social situations (high complexity with low sample size) and everyday tasks like language acquisition (I can learn a new word after hearing it just a few times, for example). Of course, vision and auditory processing break this model, but the brain need not expand its inputs into high dimensions all the time

3a) iv)

	1000×5000	5000×1000
svd	0.7524	0.6089
covariance	0.9277	0.2879

So the 1000×5000 matrices take significantly longer – this is because you're computing 5000^2 different 1000 dimensional dot products vs. 1000^2 different 5000 dimensional dot products, the first takes much longer and therefore the dimensionality matters in this respect more than the number of data points.

c) The learning without previous whitening leads to mixed images, one of which is extremely unwhitened