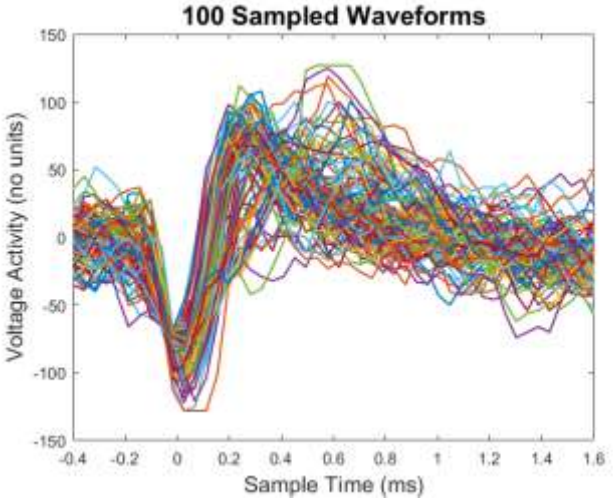
**Spike Sorting Homework**

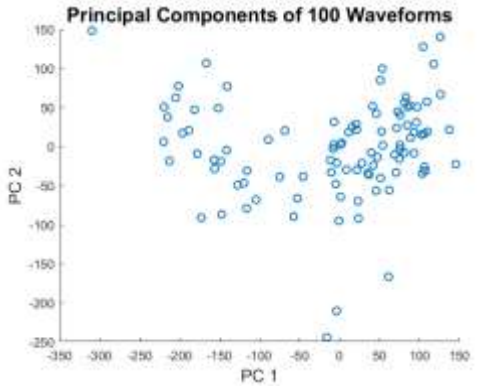
**Rory Flemming**

**BIOENG 1586 Spring 2017**

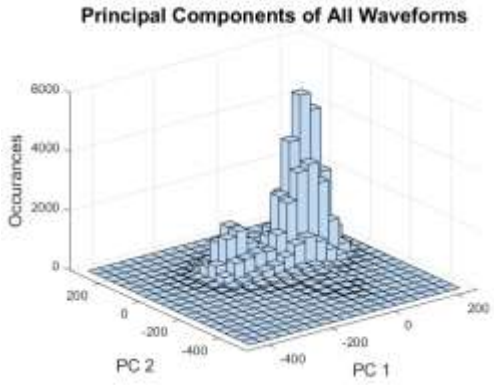



**Part 1.** Judging by eye, it is very challenging to tell how many neurons on present in the recording. There is large variance in the voltages, which makes for thick grouping. It is difficult to follow the trajectory of a single neuron in this manner. Most events seem to follow a stereotypical profile. However, it is apparent that some events have later recovery profiles (later rises than that of the main profile). Based on event amplitude, we see a range of approximately 50 units in peak activity. There may be more than one amplitude-based clusters. Pressed for an answer, I might guess that there are 2-5 distinct waveform sources in the dataset.

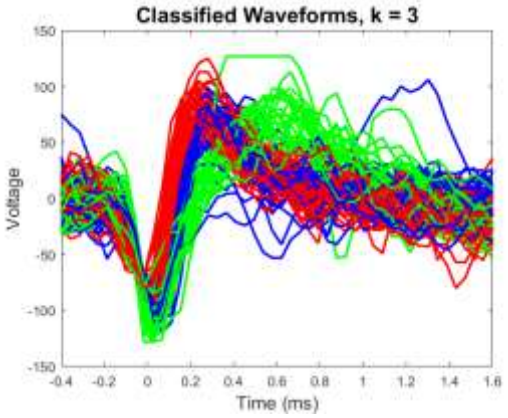
**a)**



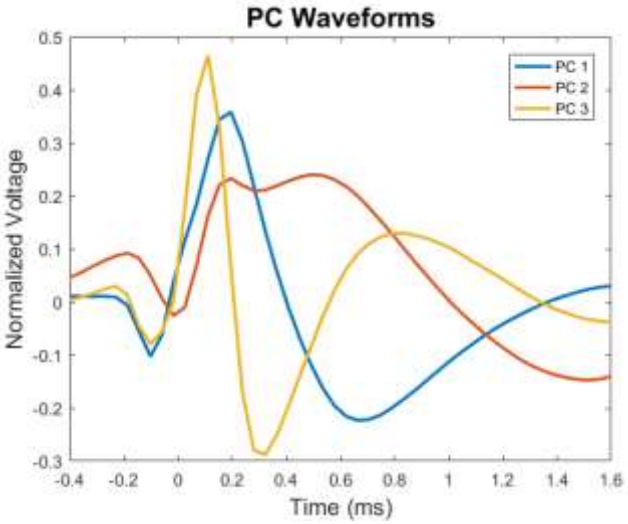

**b)**




**Part 2.** PCA of waveforms; Based on the PCA groupings, it is evident that at least two, possibly three waveform classifications may be reasonably assumed to be present within our dataset.

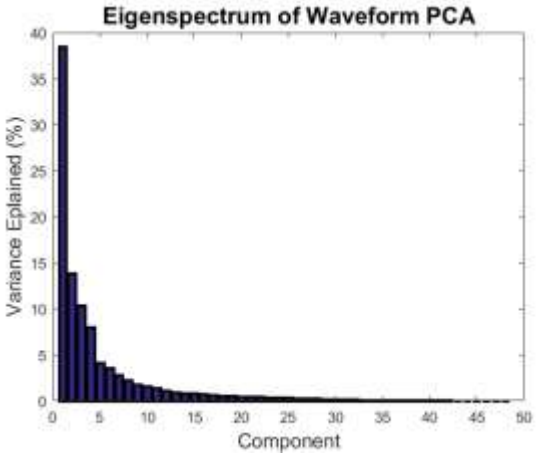



**Part 4.** K-means classification turns out reasonable groupings for 3 classes. We can see three very distinct groups especially in the rising phase. This indicates that we may likely have 3 neurons present in the recordings.

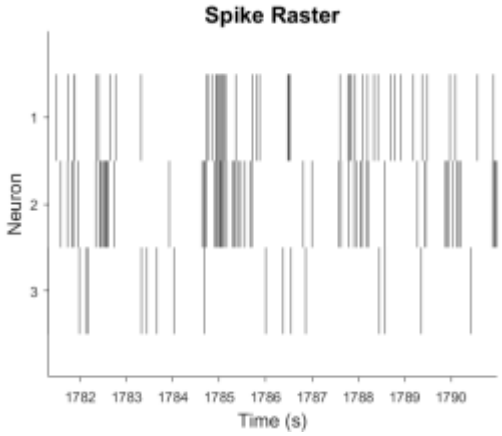



**Part 5.**

The PC waveforms appear very similar to the variable values of the neuron waveforms. This suggests that the PCA algorithm is identifying the mean values at which the most variation can be accounted for at each time point (prior components accounted for). This figure actually shows the coefficients for each component along the time axis.




**Part 6.** There are 48 dimensions in the original data. It takes 13 dimensions to account for 90% of the variance. The contribution of the dimensions is less than 5% after the 4$^{th}$, and less than 1% after the 12$^{th}$. I might suggest that the "elbow" be one of these depending on your criterion for "very little explanatory power" (5% and 1%, respectively).




**Part 7.** Spike raster of a 10s interval of the experiment. I did not color code, but I did use a conventional raster formatting and separate the neurons (code by Jeffrey Chiaou)!

**Part 8.** I actually got this far with the assumption that there were two neurons, and here I discovered that the answer was actually three! (There was an arbitrary error in my color labeling for part 4 which made it NOT informative). I would predict that ISI histograms would exhibit refractory periods, and a shape similar to the top of a whale, peaking out of water (a head and tail). This is observed when we use k = 2 and 3, however, the number of misclassifications (ISI ~ 0, before the start of the head) was high in k = 2. This is what prompted me to try k =3,4, and finally settle on 3 neurons. These histograms are consistent with well-isolated neurons, though we still do see some misclassifications. If we zoom in a little closer, we can see the actual refractory period for each neuron, since the first action potentials appear at a nonzero value of about 3ms. This links back to the inactivation of sodium channels, which require rehyperpolarization to be deinactivated.