MATLAB Assignment #5

For your previous homework you detected spikes and quantified spiking activity using peri-stimulus time histograms (PSTHs). In this homework, we will back to spike crossing and extract spike snippets to cluster different snippet waveforms. This is often called spike discrimination or spike waveform clustering. These tools are routinely used to track spiking activity from different neurons recorded by the same electrode site (or channel). A checklist is included at the end of the problem set indicating the items that will be graded. Remember to save your work as you proceed (use the save command). You can explore provided functions by using the help, type or edit commands. Accompanying questions are included at the end of each part to tie the algorithms with the neurophysiological information they intend to capture. As a guideline, answers to each question should consist of 1 to 5 sentences depending on their complexity.

Part II: Identifying and clustering different spikes

Let's step back and start again with data from one individual channel.

10. Use the raster data from above to extract snippets from the data. Use a snippet window of 1.5ms (from -0.4ms to +1.1ms). Plot a single snippet.

11. One simple way to discriminate different spikes is to look at the maximum and minimum voltage of each snippet. Plot the min vs. max of each snippet and determine whether clusters of snippets can be identified. Consider using the min and max commands for this purpose as well as a 'x' for plotting the points (e.g. plot(min,max,'x')).

12. To start moving closer towards discriminating individual spike waveforms or neurons, also called units, perform PCA on the snippet data. You can use the command svd for this purpose (see below). How many components should you consider analyzing further? This determination is often made by determining the number of components that explain variance in the data. Recall how to determine the explained variance from class. You can perform PCA using the following pseudo-code.

```
[A,e,W]=svd(snippets);
e_n=diag(e);
```

13. Recall that when using these factorization methods, snippets = A e W^T. What is A? What is e? What is W? What is " e W^T "?

14. Plot the components (i.e. eigen-vectors of A) that you think are significant and should investigated further. You may want to plot one or two more components to help your decision for number of components. Justify your selection for the number of components you think are significant.

15. Select 2 (or 3) components and plot their amplitudes (i.e. coefficients in W in this case) along a coordinated axis (see example below). Recall that these coefficients map to snippets or spikes, such that clusters in these diagrams reflect similarly shaped snippets. Is

this method helping isolate individual units? How could you separate/discriminate units based on these plots? The following commands may be helpful.

```
hist(W(:,1),100)          % plot the coefficients for the first eigen-vector
plot(W(:,1),W(:,2),'x')   % plot the map space for coeff's of eigen-vector 1 vs 2
```

16. Did your clustering leave unclustered snippets? If yes, plot unclustered snippets. Do these look like noise or real?

17. Let's go back to the PCA result and test whether we can discriminate individual units. Select 2 to 3 clusters based on their PCA coefficients either from a 1D, 2D or 3D coefficient plot. Plot the pile plot from each unit including the average unit waveform. The pseudo-code shown below shows how to keep track of discriminated units starting with raster data.

```
Recall variables "ii" and "raster" from Assignment #4,
unit1_ii = find( (W(:,1)>-0.1)&(W(:,1)<+0.1)&(W(:,2)>+0.1)&(W(:,2)<+0.3) );
unit1_snips=snippets(:,unit1_ii);
```

18. The autocorrelalogram is a useful tool to help determine whether discriminated spikes constitute a single neuron or multiple ones. Compute and plot the autocorrelalogram of all detected spikes. You can use the MATLAB command xcorr for this purpose on the raster data. Are these a single unit? The sample pseudo-code below can get you started ("t" is the time row-vector of the raw data or raster data).

```
r_xc=xcorr(raster);
t_xc=[t(end:-1:2) t(1:end)];
plot(t_xc, r_xc)
```

19. Compute and plot the autocorrelalogram of the discriminated/clustered spikes in Question #18. Could these be individual units? If you cannot resolve individual clusters, you might want to check data from other channels or just assume cluster regions.

```
unit1_raster = zeros(size(raster));
unit1_raster(ii(unit1_ii)) = 1;
unit1_xc=xcorr(unit1_raster);
```

<u>Accompanying questions</u>:

20. What is the difference between intra-cellular and extra-cellular electrophysiological recordings? Sketch a trace of their temporal shapes.
21. How can you help isolate spikes prior to data acquisition? How would you isolate individual spikes from the snippet data already acquired?
22. How could you use PCA analysis to determine whether there are there individual neurons that only respond to the visual stimulus?
23. How would you exclude snippets in a cluster that are weakly associated with the cluster?
24. How does the autocorrelalogram help us discriminate single units? What criteria is usually used for this purpose?

Accompanying questions for **extra-credit**:

a) Repeat #12-#17 in data from at least 3 or more other channels.
b) The k-means algorithm is very popular for clustering. You can easily run it on your snippets using the matlab command "kmeans" using the 'Distance' option as 'correlation'. This function returns the cluster ID for each snippet. You need to make sure the clustering results are in the snippet dimension and not in the time dimension. Justify the number of clusters selected. Select data from one channel and plot the pile-plot including the average snippet waveform for each clustered snippet in individual plots.
c) Another common use of the k-means algorithm is to cluster the PCA coefficient space using data from a channel of your choosing. In this case, you would use separate clusters by distance. Justify the number of clusters selected and plot the pile-plot of each clustered snippet along with its average snippet waveform.

Assignment Checklist

Part II: Clustering

10. Plot single snippet
11. Scatter plot of snippet min and max
12. Calculate the PCA of your snippet data. How many components should you consider?
13. Plot the PCA components (eigen-vectors) you selected and justify your selection.
14. Plot amplitudes of PCA components (eigen-values). Can you identify distinct neuronal groups?
15. Discriminate a couple of units (2-3), show their pile plots
16. Unclustered snippets? What do they look like?
17. Plot the different discriminated unit waveforms.
18. Plot the autocorrelalogram. Could it be an individual neuron?
19. The discriminated units, based on their autocorrelalogram, could these be individual neurons?
20. Answer accompanying questions 20-24 (and extra-credit a, b and/or c)