

# **ECSE 517**

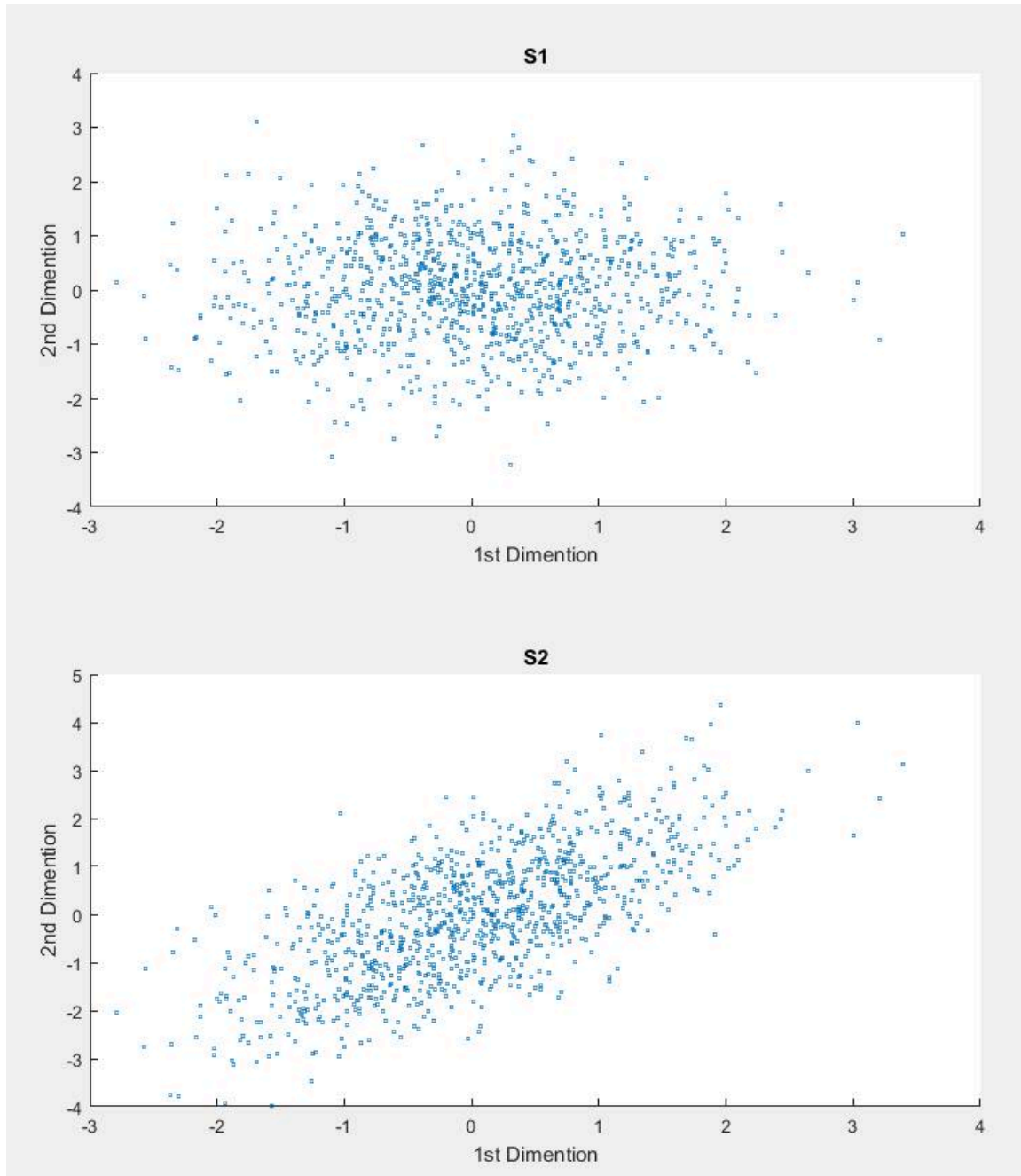
## **Neural Prosthetic Systems**

Assignment 1  
Sharhad Bashar  
I.D: 260519664

November 10, 2016

### Pre-Assignment Setup\*

Define two random data sets S1 and S2. S1 is uncorrelated, while S2 is correlated. The plots of the point are shown below:



The covariance of each of the sets are as follows. cv1 and cv2 corresponds to the covariance of S1 and S2 respectively:

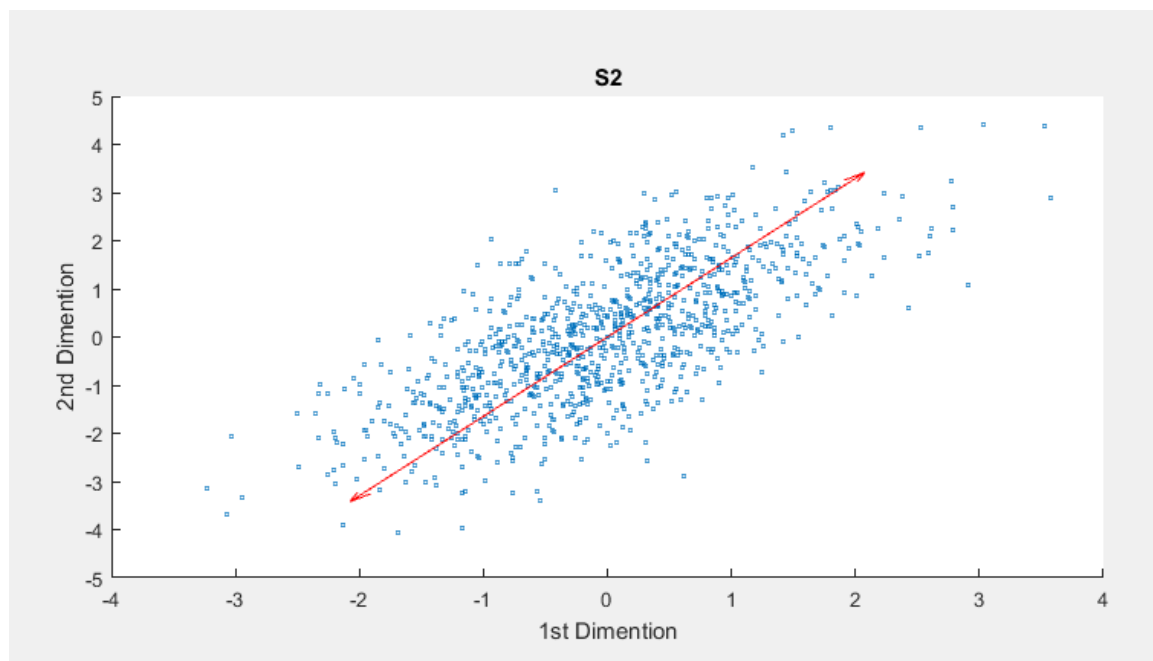
cv1 =		cv2 =	
0.9504	0.0597	0.9504	0.9695
0.0597	0.9747	0.9695	1.9807

The diagonal of the covariance matrix is the variance of each dimension since diagonal elements are the covariance of the respective dimension with itself. The off-diagonal elements are the covariance, and are equal to one another.

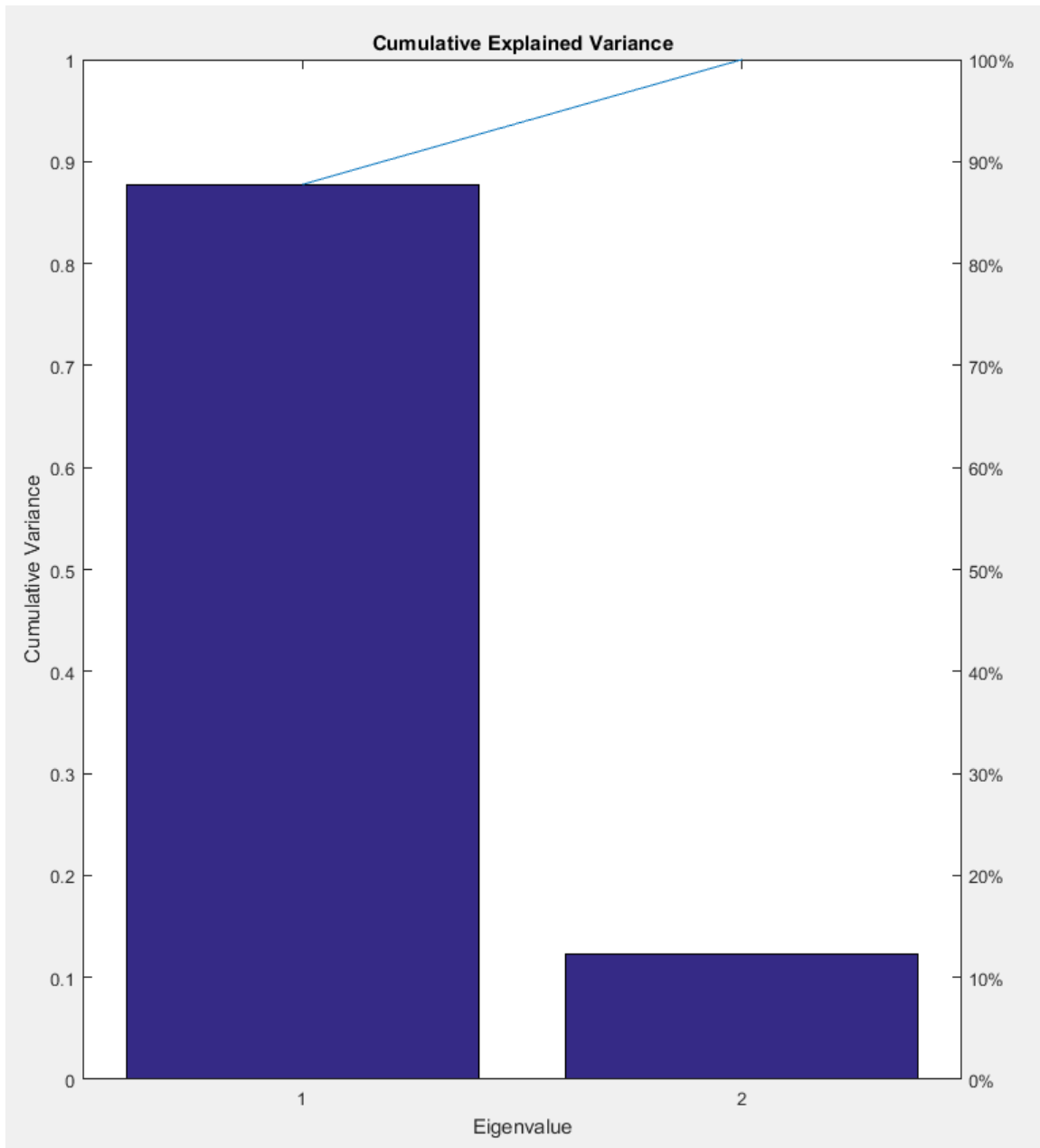
The corresponding eigenvectors, and eigenvalues of the covariance are shown below:

Evec =		Eval =	
-0.8567	0.5158	0.3900	0
0.5158	0.8567	0	2.5582

The biggest eigenvalue and the 1<sup>st</sup> principal component is 2.5582. The corresponding eigenvectors are (0.5158, 0.8567). These values are used to compute the direction of the maximum variation in S2, which is shown below with the red line:



And finally, the amount of variance explained by each principal component is shown below:



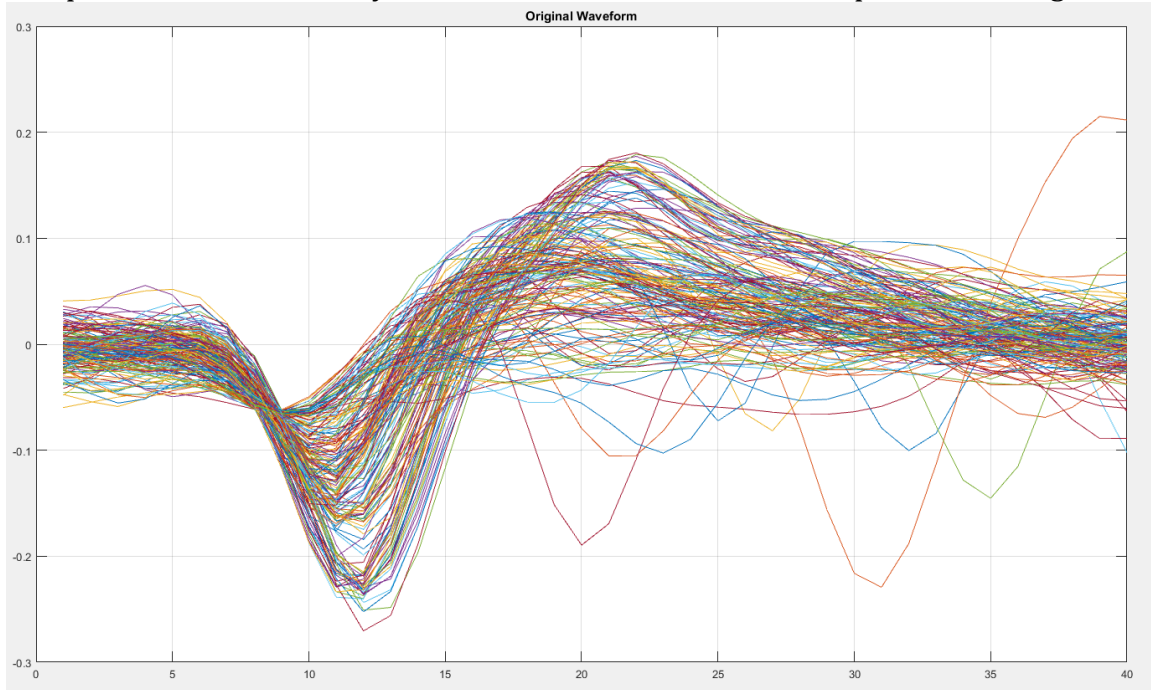
The first PC captures  $\left[ \frac{2.5582}{(2.5582 + 0.39)} * 100\% \right]$  or 86.7% of the variation, and the second PC captures the rest 13.2%

*\*Code for pre assignment is included in the Appendix*

## Part A: Dimensionality Reduction\*

### 1. Clean up data for spike sorting

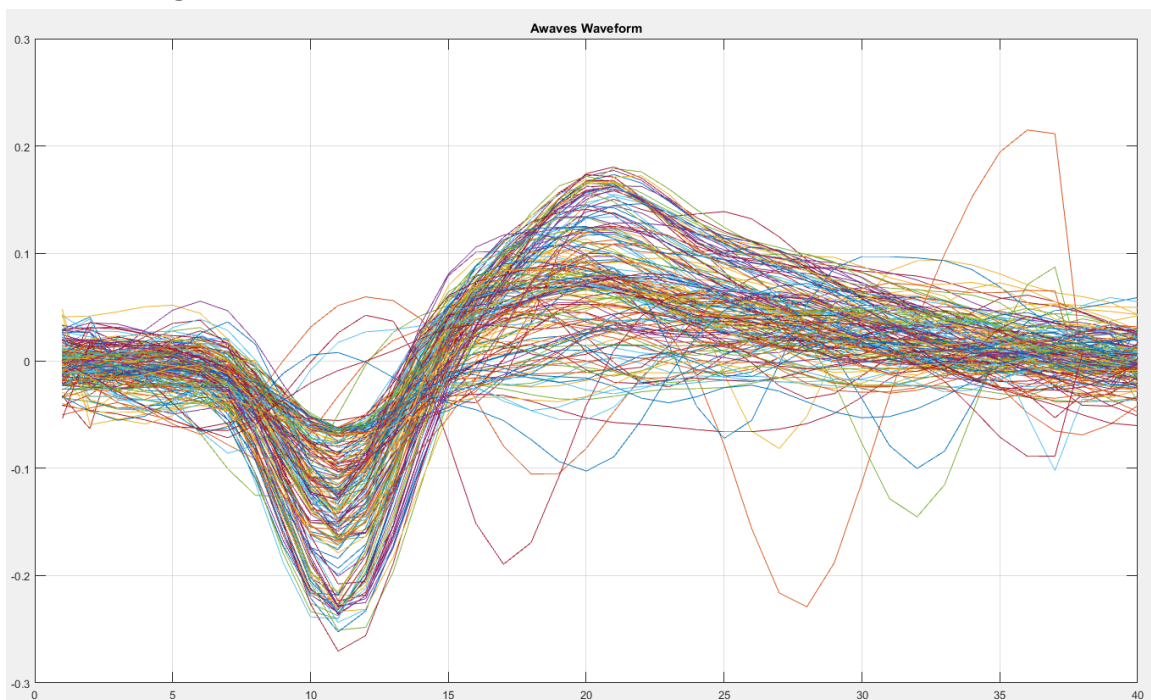
The plot below shows every 100<sup>th</sup> waveform in raw form. The plots are unaligned:



The following steps were taken in order to align them:

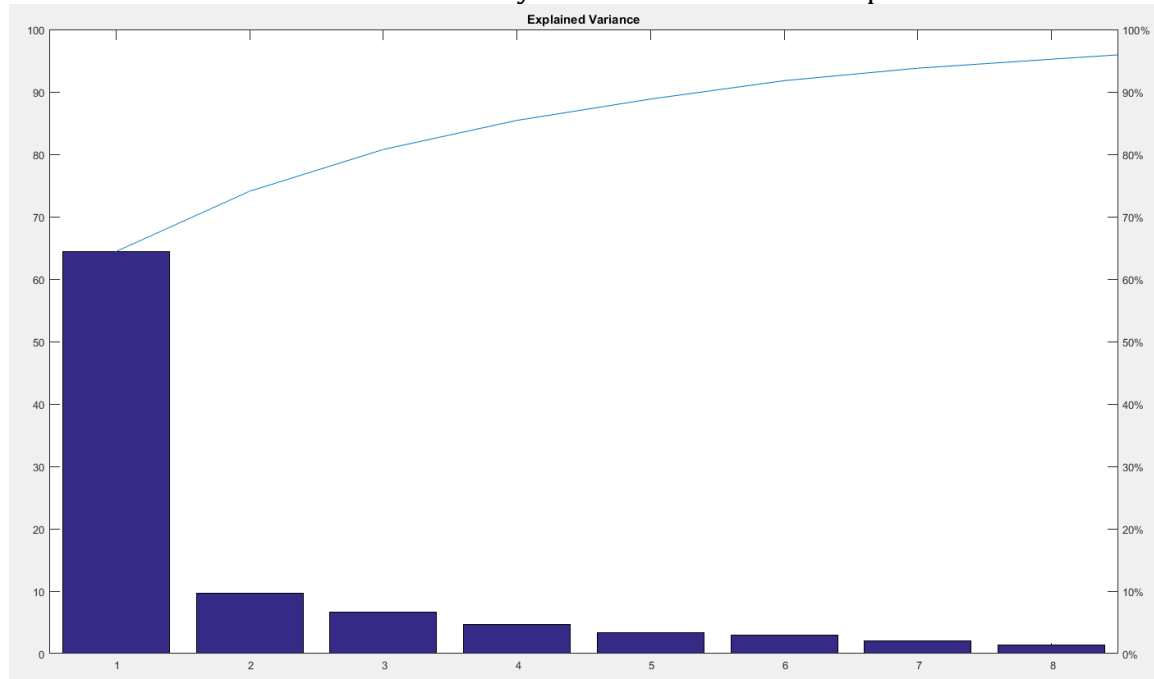
- First the time where the minimum occurred for each of the waveforms were calculated.
- The mode of these time sets were calculated, and was found to be at  $t = 11$ .
- If a waveform reached its minimum at a time value within 3 steps of the mode ( $8 \leq t \leq 14$ ), it was shifted the appropriate amount to 11, and for other wave forms, they were shifted by 3 units towards 11.

The final Aligned waveform looks as follows:



## 2. Run PCA on Awaves

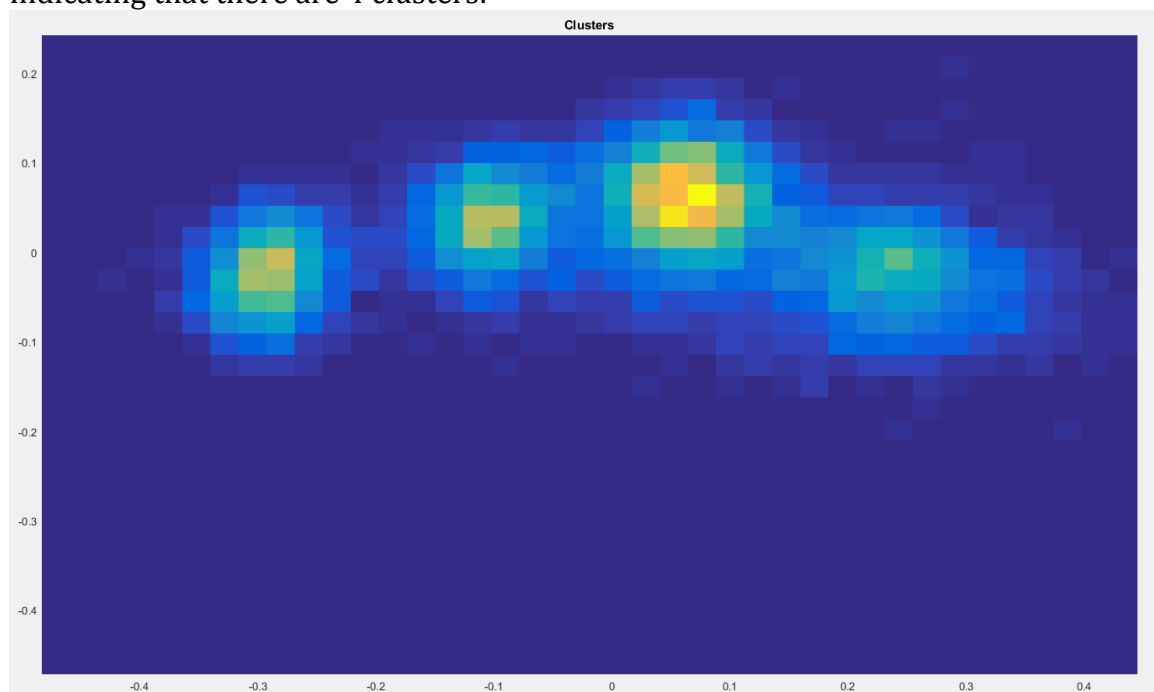
The cumulative variance accounted by each PC is shown in the plot below:



90% of the variance is accounted for in the first 6 PC's. 8 PC's account for 95% of the variance. And less than 0.1% of the PC's are accounted by 23 PC's.

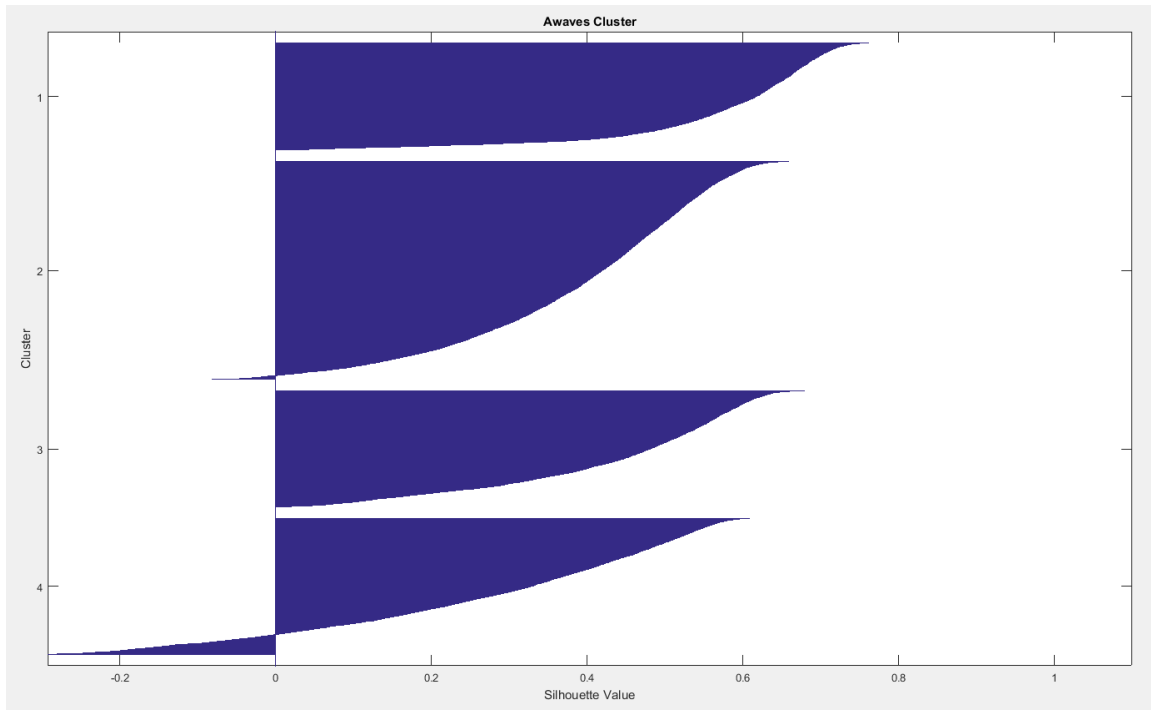
## 3. Use a scatterplot to plot PC1 vs PC2

The scatter plot below compares the two PC's. We can see 4 yellow groupings, indicating that there are 4 clusters:

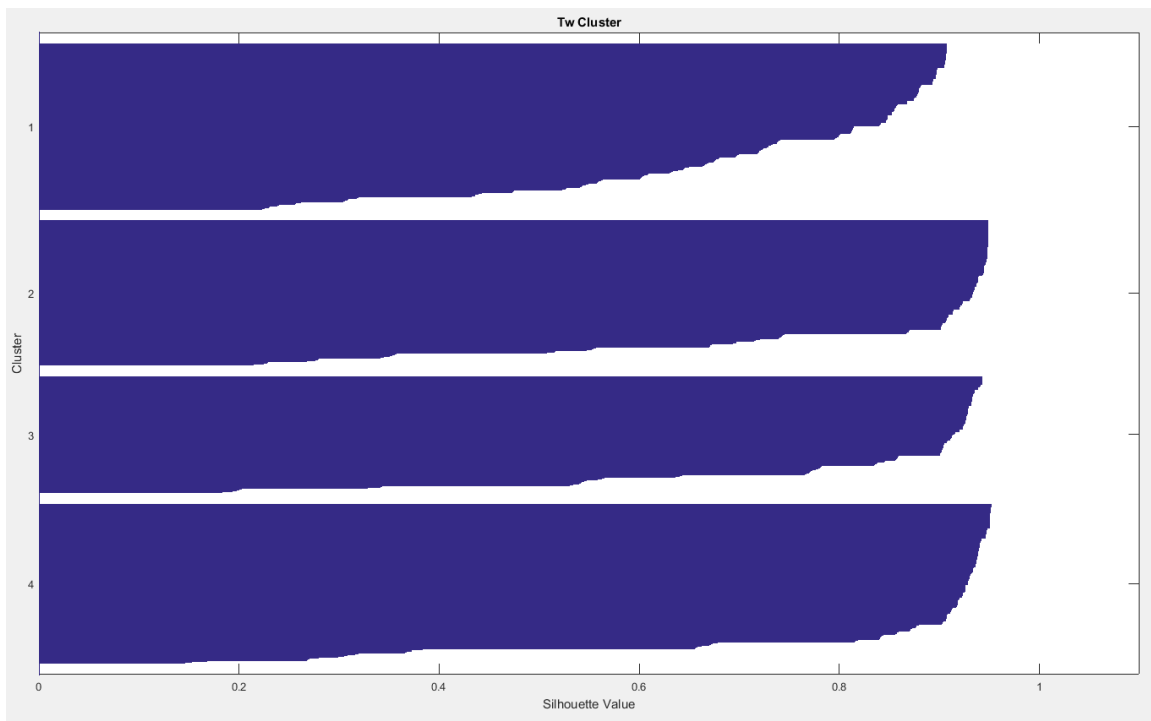


#### 4. Use Kmeans to cluster the data

There were 4 clusters, so kmeans returned membership indices ranging from 1 to 4. The waveform were grouped accordingly, and produced the following plot:



Plotting the spike times in data.TW gave the following plot:



As we can see from the first chart, there are indeed 4 main groups. The silhouette values range from -1 to +1. +1 indicates points were distinct from other clusters. 0 means points weren't in any particular cluster. And -1 means that a point is in the wrong cluster.

In the TW data, there are 4 clusters as well, and we can clearly see that the clusters are well separated, and that most of the values are closer to 1. This indicates that the points are well organized and separated from the other clusters.

*\*Code for Part A is included in the appendix*



# Appendix

---

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Course:    ECSE 517
%Assignment: Assignment 1
%Author:    Sharhad Bashar
%File name: assign_1_startup.m
%Description: This file contains all the pre-assignment requirements
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

```
%Random data set S1
%1000 x 2 points, mean = 0, variance = 1
S1 = randn(1000,2);
```

```
%Set S2:
%First coloumn is x axis, second coloumn is y axis
S2(1:1000,1) = S1(:,1);
S2(1:1000,2) = S1(:,1) + randn(1000,1);
```

```
%Calculate the covariance of the two sets
cv1 = cov(S1);
cv2 = cov(S2);
```

```
%Eigenvector of covariance
[Evec,Eval] = eig(cv2);
```

```
[BigEvec,BigEval]= eigs(cv2,1);
BigEvec = BigEvec * (-1); % becuase BigEvec was returning negative values of the
largest row
```

```
figure;
subplot(2,1,1);
scatter(S1(:,1),S1(:,2),4);
title('S1');
xlabel('1st Dimention');
ylabel('2nd Dimention');
subplot(2,1,2);
scatter(S2(:,1),S2(:,2),4);
title('S2');
xlabel('1st Dimention');
ylabel('2nd Dimention');
hold on;
quiver(0,0,BigEvec(1,1),BigEvec(2,1), 4,'color', 'red');
hold on;
```

```

quiver(0,0,-BigEvec(1,1),-BigEvec(2,1), 4,'color', 'red');
hold off;
SortEval = sort(diag(Eval),'descend');
PCnorm = zeros(size(SortEval));
for i = 1:size(SortEval);
    PCnorm(i)=SortEval(i)/sum(SortEval);
end

%Plot the data bases
figure;
pareto(PCnorm);
title('Cumulative Explained Variance');
xlabel('Eigenvalue');
ylabel('Cumulative Variance');

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Course:    ECSE 517
%Assignment: Assignment 1
%Author:    Sharhad Bashar
%File name: assign_1.m
%Description: This file contains the code for Dimensionally Reduction
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

Waves = Data1.Waves;
data = Waves(1:end,:);
Awaves = Waves;
dataSize = size(data,2);

%Part 1
[minVal,Index] = min(data); % Gets the min of the data set, as well as the location of
the minimum
minIndex = mode(Index); % stores them in a data
modeIndex = mode(minIndex);
for i = 1:dataSize
    displacement = Index(i) - modeIndex;
    if (abs(displacement) > 3)
        displacement = 3;
    end
    if (displacement ~= 0 )
        Awaves(i,:) = circshift(Awaves(i,:),[0 -displacement]);
    end
end

figure;
plot (Waves(1:100:end,:));
grid on;

```

```
title('Original Waveform');
figure;
plot (Awaves(1:100:end,:));
grid on;
title('Awaves Waveform');
```

```
%Part 2
[coeff,score,latent,tsquared,explained,mu] = pca(Awaves);
figure;
pareto (explained);
title('Explained Variance');
```

```
%Part 3
edg = prctile(score(:,1:2),[0.1,99.9]);
De = min(diff(edg))/30;
e1{1} = (edg(1,1):De:edg(2,1));
e1{2} = (edg(1,2):De:edg(2,2));
hst = hist3(score(:,1:2),e1);
Shand = pcolor(e1{1},e1{2},hst');
title('Clusters');
set(Shand,'LineStyle','none');
```

```
%Part 4
idxAw = kmeans(Awaves,4);
figure;
[silh,clust] = silhouette(Awaves,idxAw);
title('Awaves Cluster');
xlabel ('Silhouette Value');
ylabel ('Cluster');
```

```
Tw = Data1.Tw;
idxTw = kmeans(Tw,4);
figure;
[silh,clust] = silhouette(Tw,idxTw);
title('Tw Cluster');
xlabel ('Silhouette Value');
ylabel ('Cluster');
```