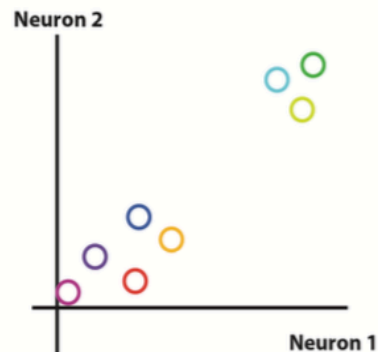


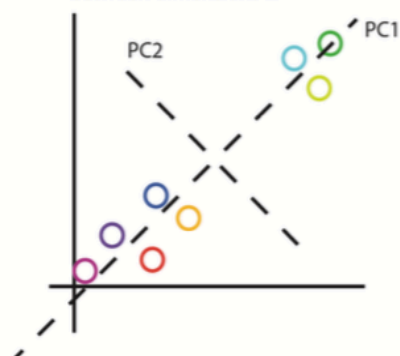
A.

Original Data is highly correlated



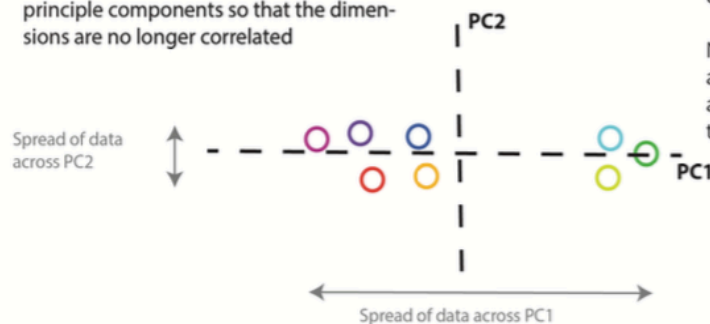
B.

PCA first finds the lines, or principal components that best 'capture' correlations between dimensions ...



C.

... and then replots the data along these principle components so that the dimensions are no longer correlated



PCs are sorted by the information still contained along that dimension.

Notice that the data is most spread out across PC1, therefore it tells you the most about the network's state at any given time

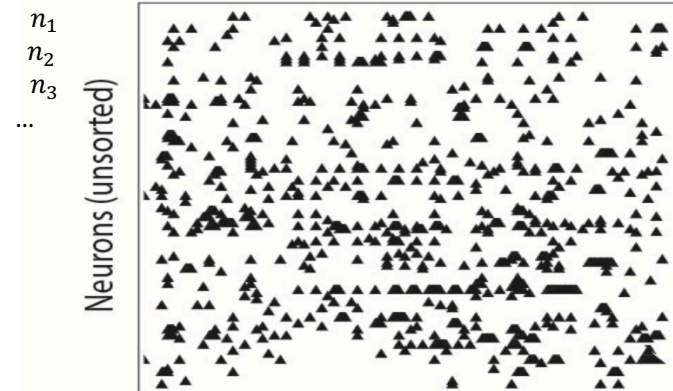
D.

After PCA the data is effectively captured by one dimension: PC1

Practically no information is lost by removing PC2 - thus we have **reduced the dimensions** of our dataset

E.

This could be tens of thousands of neurons



PCA (from time perspective)

$$N_1(t) = a * n_1(t) + b * n_6(t) + c * n_{2000}(t) \dots$$

$$N_2(t) = d * n_1(t) + e * n_{10}(t) + f * n_{12}(t) \dots$$

 $N_1(t)$ and $N_2(t)$ orthogonal. $[a, b, c, \dots]$ and $[d, e, \dots]$ orthogonal.

F.

TPC of neurons activity with time from 9 slices averaged from 3 fishes

