# Spike Waveform PCA pipeline

1. wavePCA.m
   * Specify the experiment/probe you want to spike sort and the proportion of spikes from that experiment you want to use for calculating PCs.
   * Inputs:
     1. animalID = animal ID (string, ex: ‘febe0’)
     2. unitID = unit ID (string, ex: ‘000’)
     3. expID = experiment ID (string, ex: ‘000’)
     4. probeID = which probe in this experiment (number)
     5. nJobs = number of jobs SpikeFiles are broken up into (number)
     6. propSpks = proportion of spikes used to calculate PCs (number)
        + for experiments with many spikes detected, reducing the proportion can greatly speed things up while still producing very similar PCs.
     7. dataFold = path to folder storing data to be processed (string)
        + should contain folders for each animal ID > folders for each individual experiment
   * Output:
     1. File: fexx0\_000\_000\_pX\_pca.mat
     2. Contains structure ‘PC’ with fields:
        + fileName
        + coeff = pxp matrix corresponding to the weights for each principal component (column)
        + latent = px1 vector corresponding to the variance explained by each PC
        + explained = px1 vector corresponding to the percentage of total variance explained by each PC
        + score = nxp matrix of ‘wvfrms’ projected into PC space (each row is a waveform, each column is a PC)
        + wvfrms = nxp data matrix of waveforms that input to pca to get PCs (n = number of waveforms, p = number of samples per waveform)
        + mu = mean(wvfrms); 1xp vector that is the average waveform
        + propSpks = proportion of spikes from this experiment used to generate PCs
   * to project waveforms into PC space:

score = (wvfrms-mu)\*coeff

1. spkWvPC.m
   * uses the info in structure ‘PC’ generated by wavePCA to project spikes into PC space and save projection of first 3 PCs for each spike in SpikeFiles
   * deals with one job at a time (run in parfor loop over all jobs)
   * Inputs (see explanations above):
     1. animalID
     2. unitID
     3. expID
     4. probeID
     5. job
     6. dataFold
   * Output:
     1. File: fexx0\_000\_000\_jX\_pX\_spkinfo.mat
        + Contains ‘spk’ structure updated with PC1Det, PC2Det, and PC3Det (the projections of each spike onto first 3 principal components)
2. alignWvs.m
   * function used in wavePCA to align waveforms in data matrix to their minimums