BIOL 450 / IBS 534 HW1 Spike Train Analysis

Name: Ruorong Qi
Date: Jan 2025

Part I

1. Use the Matlab built-in function randperm to plot STA curves using N =100, 500, 1,000, 5,000, 10,000, and 25,000 spikes. Also calculate using all 53,601 spikes. Make sure to save the function's output, as you will need the saved STA curves for subsequent questions.

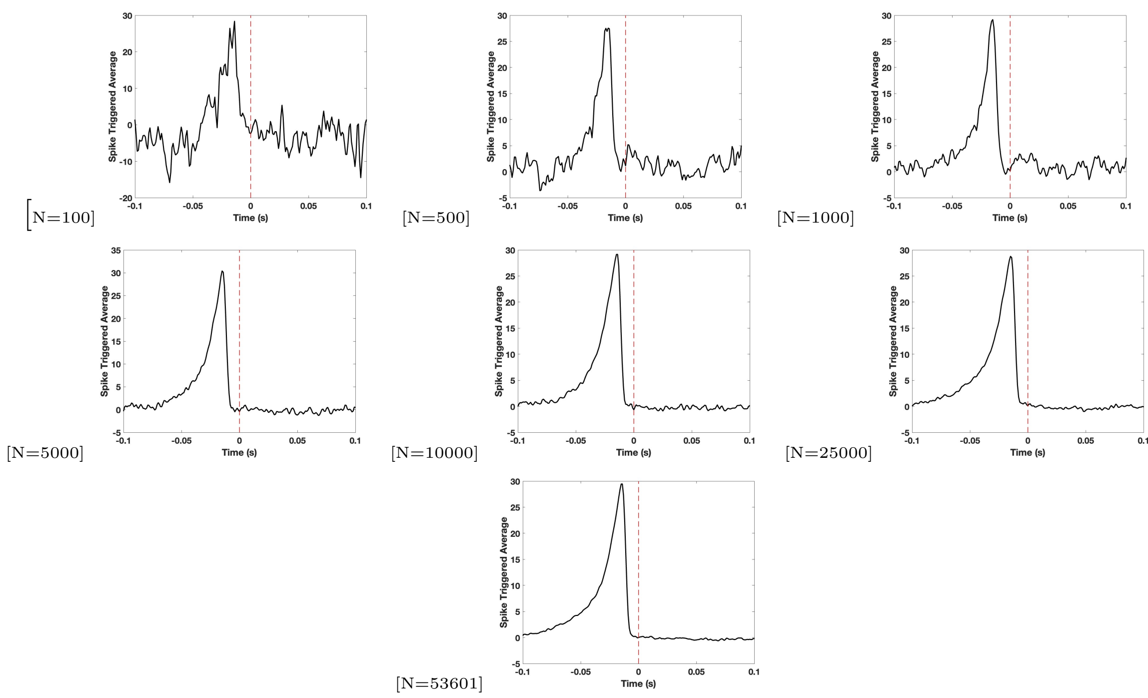

Figure 1: STA curve

2. What happens to the curve as N becomes larger?
   As N becomes larger, STA curve becomes smoother (less fluctuation).

3. Now let's try to quantify how the average changes as N becomes large. If $s_k(t)$ is the STA curve using k spikes, calculate

$$d(k) = \sqrt{\sum_t [s_k(t) - s_{53601}]^2}$$

   for each of the values of k you previously calculated. Plot d(k) vs. k on a log-log plot (loglog in Matlab).

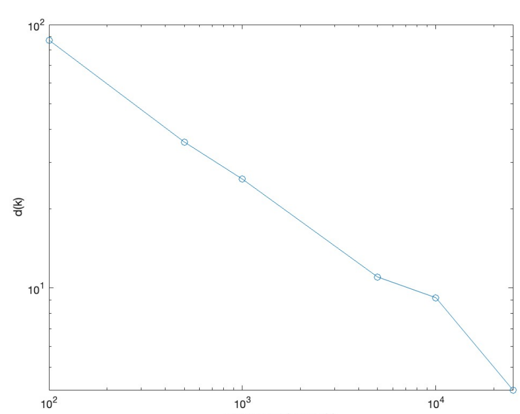

Figure 2: d(k) vs. k

4. If a function is of the form $y(x) = Ax^b$, then the slope on a log-log plot provides an estimate of b. What is (roughly) the slope of the line you see? What does this say about how the accuracy of the measured STA improves as one watches the data for increasingly long periods of time?
   The Slope of this log-log plot is roughly -0.5, which means STA improves as we watch it for long periods of time.

Part II

1. The linear potion of these types of models builds-off of the STA (for all spikes) that you found in the previous part of the homework set. We will call this STA s(t). To determine the score, or the similarity of match to the filter, we can compute the convolution between the data set, x(t), and s(t) using

$$X(t) = \sum_{t'} x(t - t')s(t')dt$$

   where $X(t)$ is the value of the convolution at time t and dt is the sampling time of the data (here, .001 seconds). Use convolutionOutput = convolveDataWithSTA(data,STA,dt) to calculate this sum for all time points. Plot the found convolution for the first second of the data set.
   The sum for all time points is -3.7411e+04.

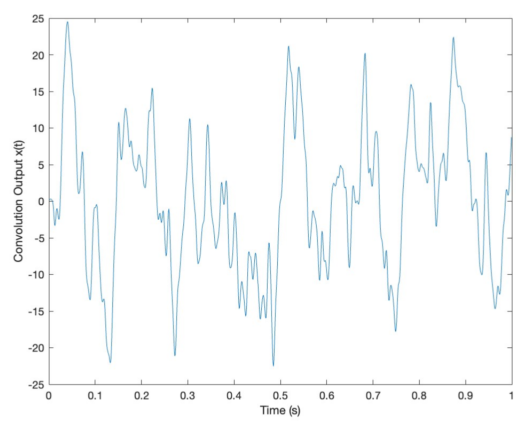

Figure 3: Convolution Output in First Second

2. Use plotSpikesOnData(H1 times,convolutionOutput,H1 spikes,startTime,endTime) to plot the first second of the convolution superimposed with the locations of the spikes (the red vertical lines). Here, startTime is 0 and endTime is 1. What happens to the convolution function when a lot of spikes occur?
   The convolution function increase drastically when a lot of spikes occur.

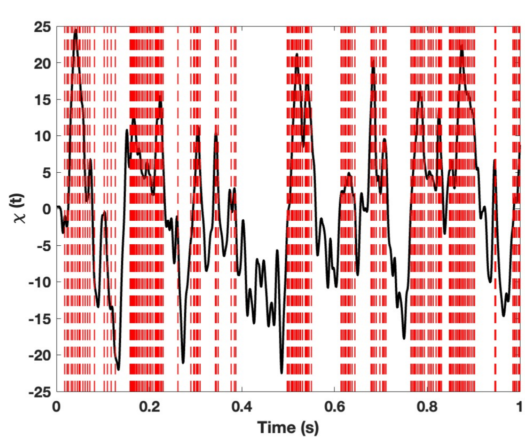

Figure 4: Spike on Convolution Output in First Second

3. Use the hist function in Matlab to make a histogram of the values of $X(t)$ you previously calculated. Use 20 bins, and make sure to save the output to your workspace, as we will use this in a moment.

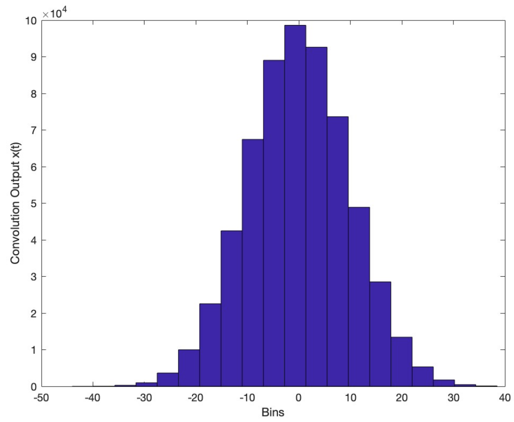

Figure 5: Binned Convolution Output Distribution

4. Now,we would like to map these convolution scores,$X(t)$ on to a probability of firing, $p(spike|X)$. In other words, we want to find a function, $f(x)$ that takes the convolution value to something between zero and one, but we will let the data tell us what that function is! To do this, we will calculate the conditional probability of observing a spike, given that a value of $X(t)$ is observed (i.e. the number of spikes when $X(t) \approx y$ divided by the total number of time points where $X(t) \approx y$). The loops needed to achieve this are a little tricky, so I have provided code to find this: [binLocations,pSpike,numSpikes] = findSpikingProbabilitiesFromConvolution(convData,spikes,numBins), where numBins is the number of bins (use 20), convData is the convolution from before, and spikes are the spike locations. Display this plot.

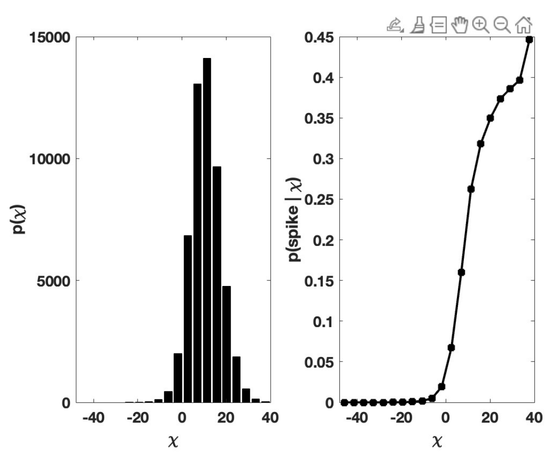

Figure 6: (left) Number of Spikes vs. Binned Convolution Output;
(right) $p(spike|X)$ vs. Binned Convolution Output

5. What does the curve on the right ($p(spike|X)$) vs. X) look like? Does this agree with your intuition from the plot in question 2 of part II?
As convolution increase, the probability of spiking given that convolution increase sigmoidally. This agree with my intuition for Q2.2 that spiking probability increase with increasing convolution.

6. Given a new stimulus presentation to the same neuron, describe how would you predict the location of spikes using the STA and the curve you derived in this homework set?
Firstly, compute the STA based on new stimulus and spikes, the new stimulus is then convolved with STA to compute convolution output $X(t)$, which reflect how similar is the new stimulus to a typical spike triggering event; Then, the probability of spiking event is computed by mapping the convolution output $X(t)$ to probability curve $p(spike|X)$; Finally, predict the location of spikes based on the probability curve.