Project -3: Computational Neuroscience

Team 2*: Vedic Partap, Shounak Sural, Prashanthi Silla, Param Kishor, Ravi Sheoran

Topics

Encoding and Decoding

Pre-Processing

Divide the total 20 sec of data in the training and testing set. First 15 seconds of data for estimating the model/receptive field and the last 5 seconds of data for prediction purposes.

Questions

Q1: Stimulus Nature

A1: From the first observation, the output of the auto-correlation of the stimulus resembles with the Dirac Delta Function. The autocorrelation function is normalized. The output is Figure -1

MATLAB: `xcorr(Stimulus, max_delay,'normalized');`

Q2: PSTH and mean firing rate

A2: The data is divided into training and testing datasets. The bin size is 1 ms. There is a total of 15000 (15sec/1ms) bins in the training set and 5000 (5sec/1ms) bins in the testing set. Figure-2 for the training set and Figure-3 for the testing set.

Q3: Poisson or Non-Poisson

A3: For a Poisson process the mean and variance is equal.

$$P\{N(t) = n\} = \frac{(\lambda t)^n}{n!} e^{-\lambda t}.$$

The mean and variance are λ.

By plotting the variance vs. mean, we can check if it a Poisson process if data points lie on $y = x$

The mean and variance are found for the distribution of spike count for each bin size. The bin size varies from 10 ms, 20 ms, 50 ms,100 ms, 200 ms, and 500 ms.

As we can see, the Poisson nature of the distribution reduces as the bins size increases.

* Contributions at Page 12

Figures: 4-9 show the mean vs. variance for different bin sizes. In later figures the spread across y=x increases

Q4: Spike Triggered Average (Correction for non-Gaussian)

A4: Assumption: Spike only depends on the previous 100 ms (window size) of the stimulus.

The Spike triggered average is found for all the four neurons. The number of spikes arriving before 15 sec (training set) is found, and cumulative STA is calculated. It is averaged out.

Correction for Non-Gaussian

```
cssinv=inv(corr_mat);
```

The correction term is found by multiplying a corrective factor CSS^{-1}, which is inverse of the auto-correlation matrix. Since results got worse with the use of CSS^{-1} as input was already Gaussian, the original waveform is considered for the rest of the project

Figure-10 (without corrections) is also not a Gaussian distribution, but Figure-11 is not an improvement over Figure-10.

Q5: Determining the output non-linearity

A5: Method:

1. Predict the λ, parameter for the point process distribution.
2. Plot estimated and observed value.
3. Try to fit a function from these values.

On fitting Sigmoid function using a freely-available sigmoid-fitting code, the results were worse mainly due to the neuron one which has a slight negative slope in linear fit and has huge errors for sigmoidal fit as shown in Figure-12

Q6: Prediction performance and pruning of filter parameters

A6: The problem when fitting the function is over-fitting. If the number of correct data point increase then the fitting become better. But in the presence of noisy data points out model starts fitting these points.
To avoid that, we test our model on the testing set, and if the difference between training and the testing accuracy is significant, then our model has over-fitted the training set.

Metric used to find the compare the predictions is Square of correlation between the predicted and observed distribution.

Figure -13 shows the linear fit on the data. It can be observed that for neuron 2,3 it is a good fit but not for neuron 1,4

Figure 14,15 shows the measured vs. predicted lambda values for all the four neurons.

Figure 16 shows the r^2 values.

Q7 B: Victor and Purpura (VP) Spike Distance Metric

A7 B: It is the distance between two spike trains in terms of the minimum cost of transforming one spike train into the other by means of just three basic operations: spike insertion (cost 1), spike deletion (cost 1) and shifting a spike by some interval Δt (cost $q|\Delta t|$).

It is similar to the edit distance problem in computer science.

Optimal Timescales (1/q): It can be seen from Figure-17 that the information is the maximum of the intermediate values of q.

Figures

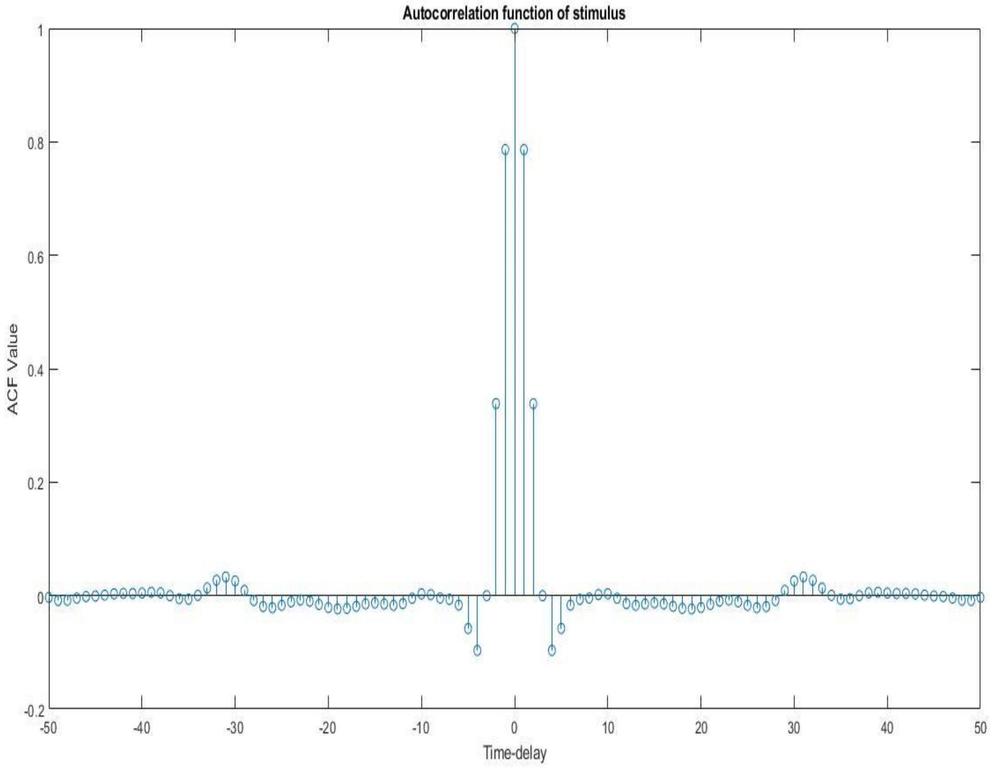

Figure 1: Normalized Auto-correlation of the stimulus

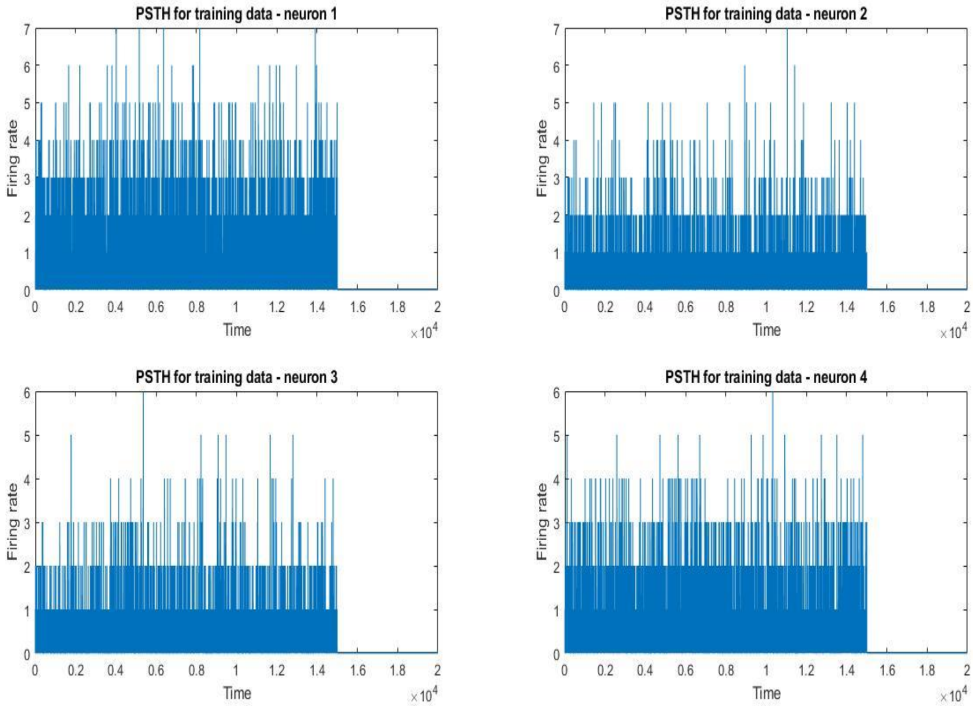

Figure 2: PSTH for the training set (15sec)

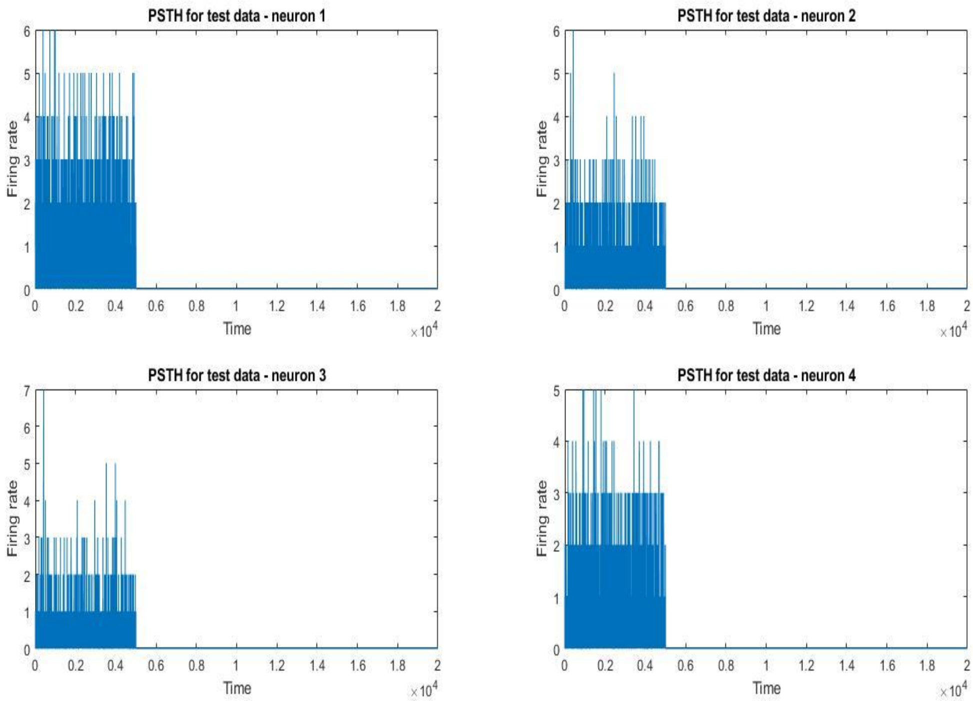

Figure 3: PSTH for the testing set (5sec)

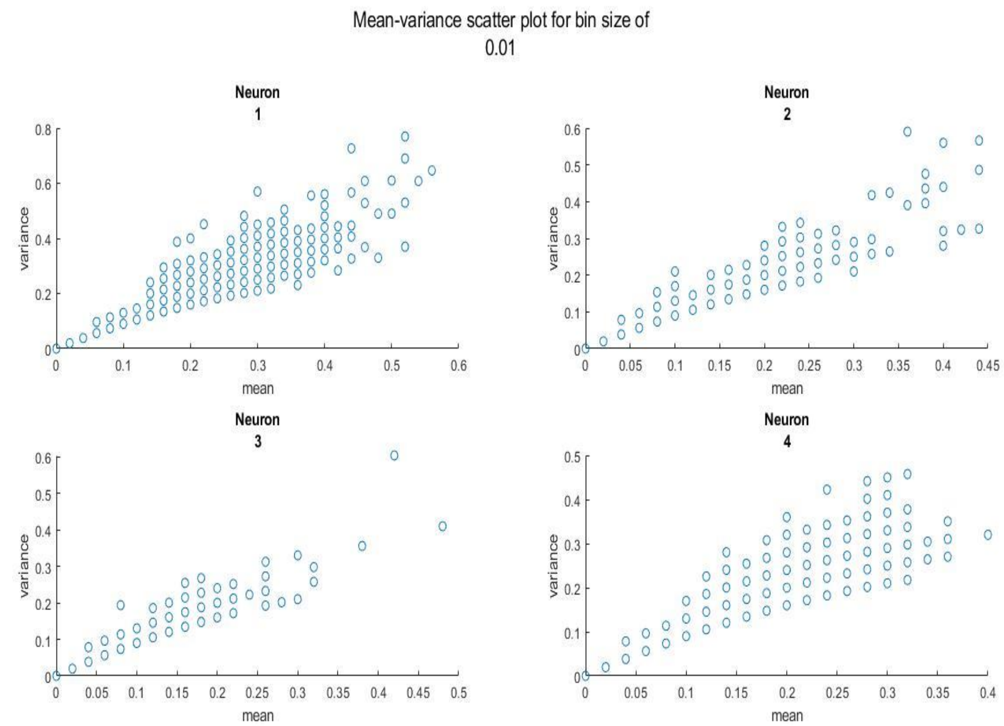

Figure 4: Mean vs Variance bin size = 0.01

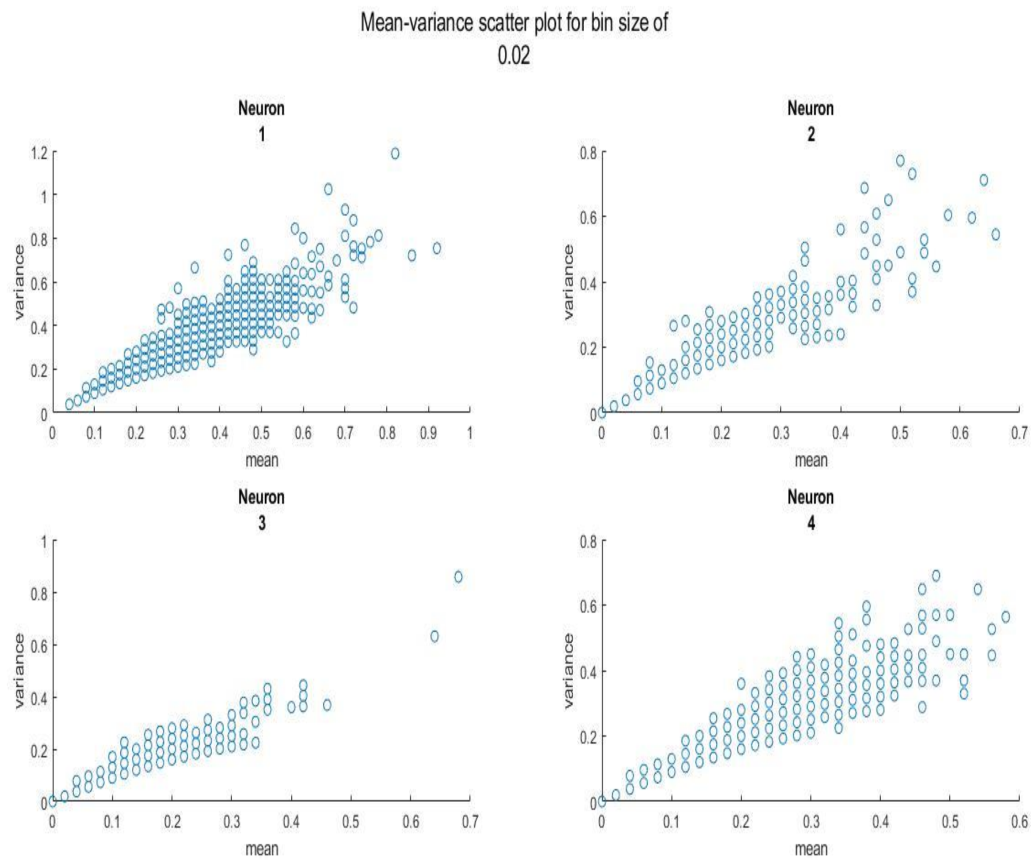

Figure 5: Mean vs Variance bin size = 0.02

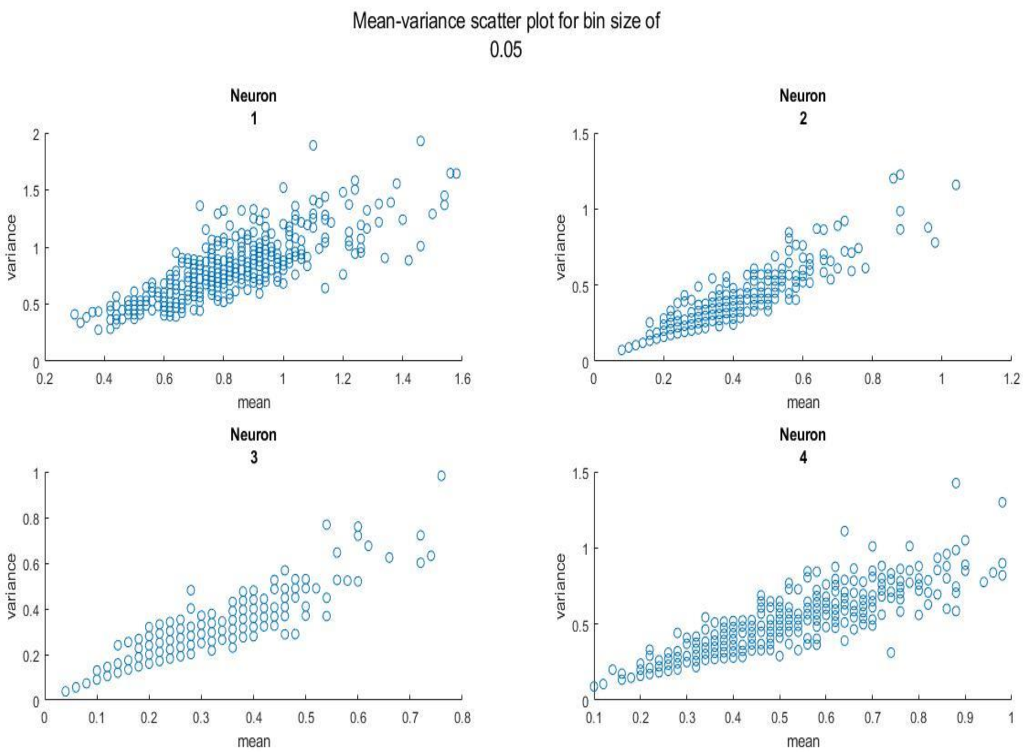

Figure 6: Mean vs Variance bin size = 0.05

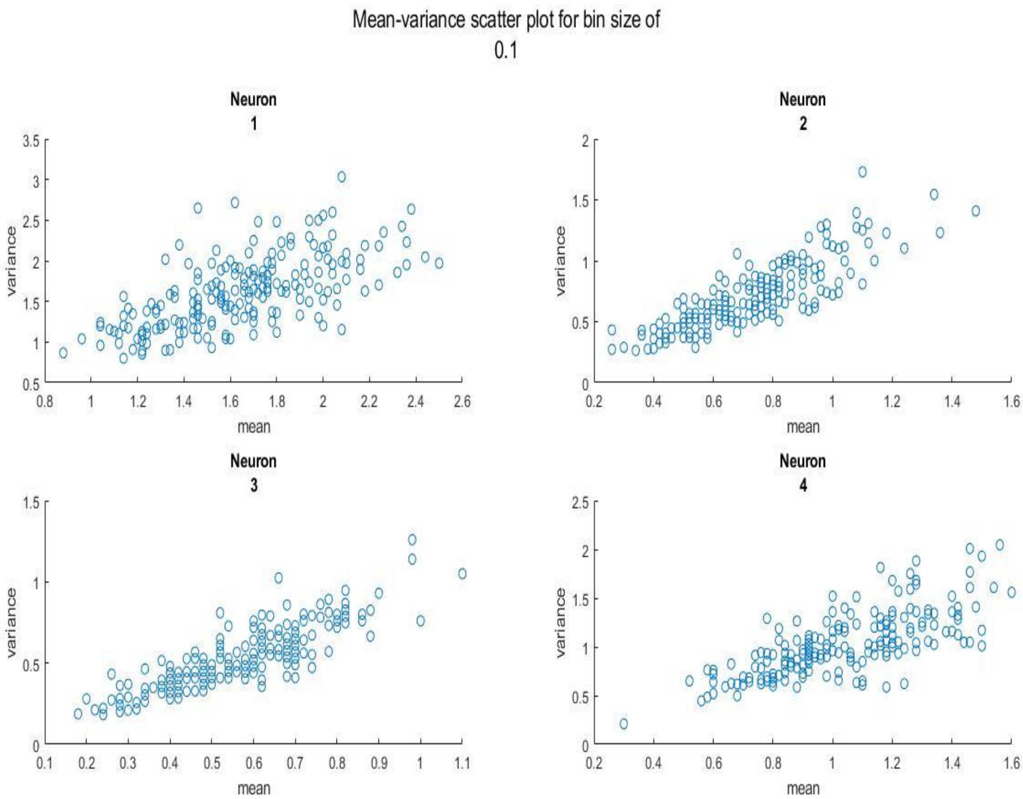

Figure 7: Mean and Variance bin size = 0.1

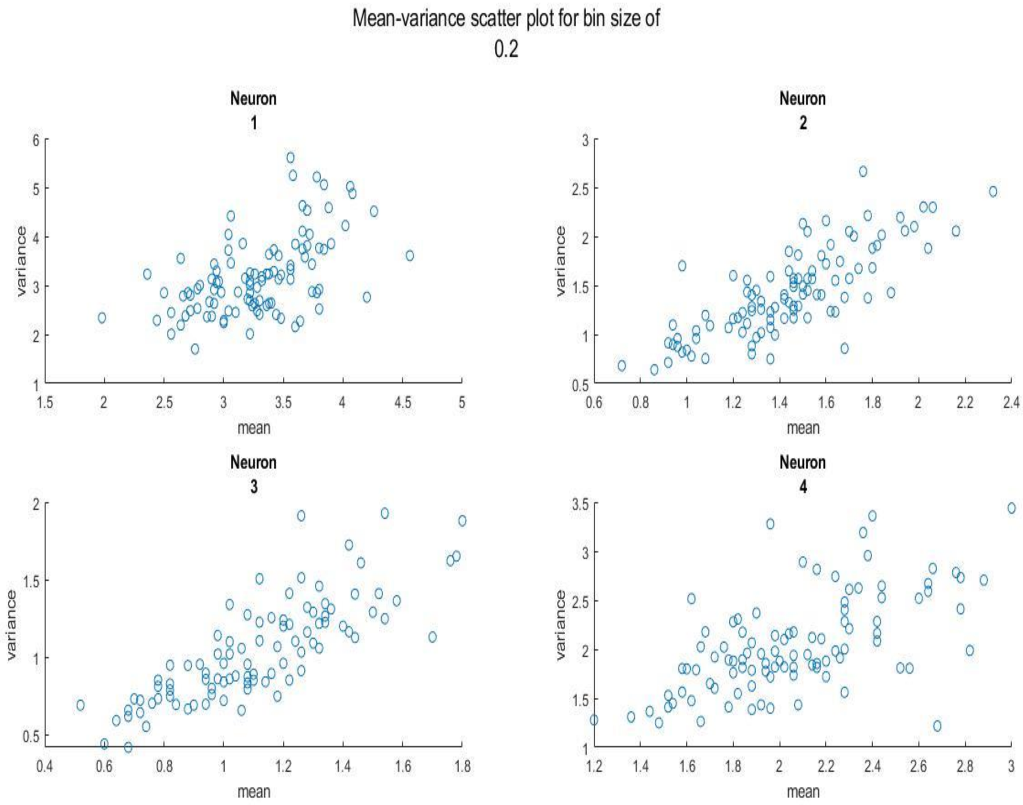

Figure 8: Mean and Variance bin size = 0.2

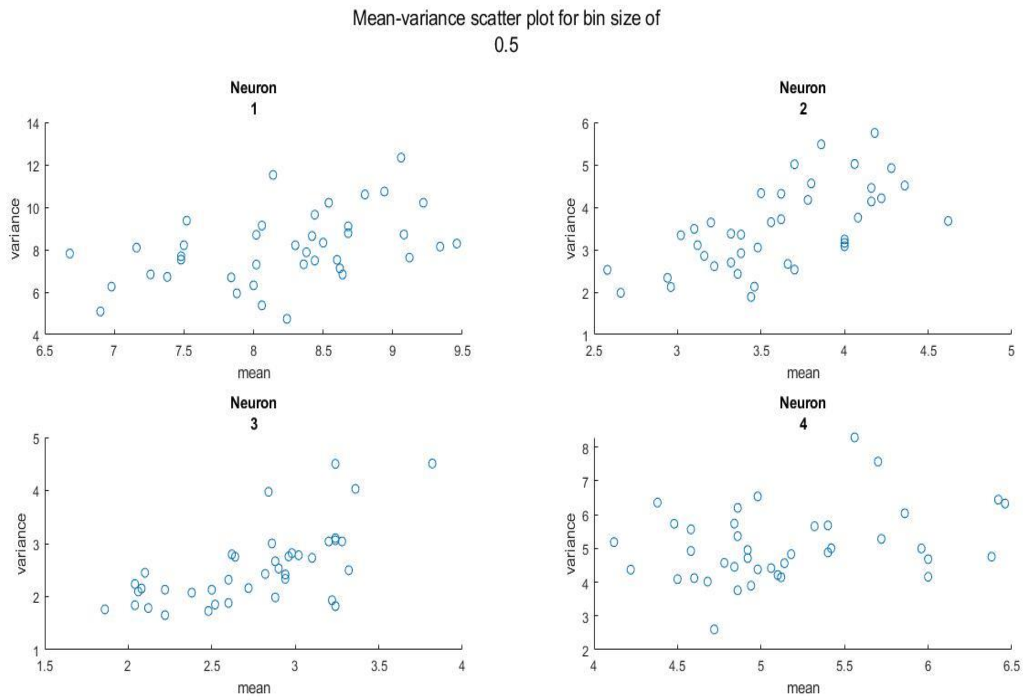

Figure 9: Mean and Variance bin size = 0.5

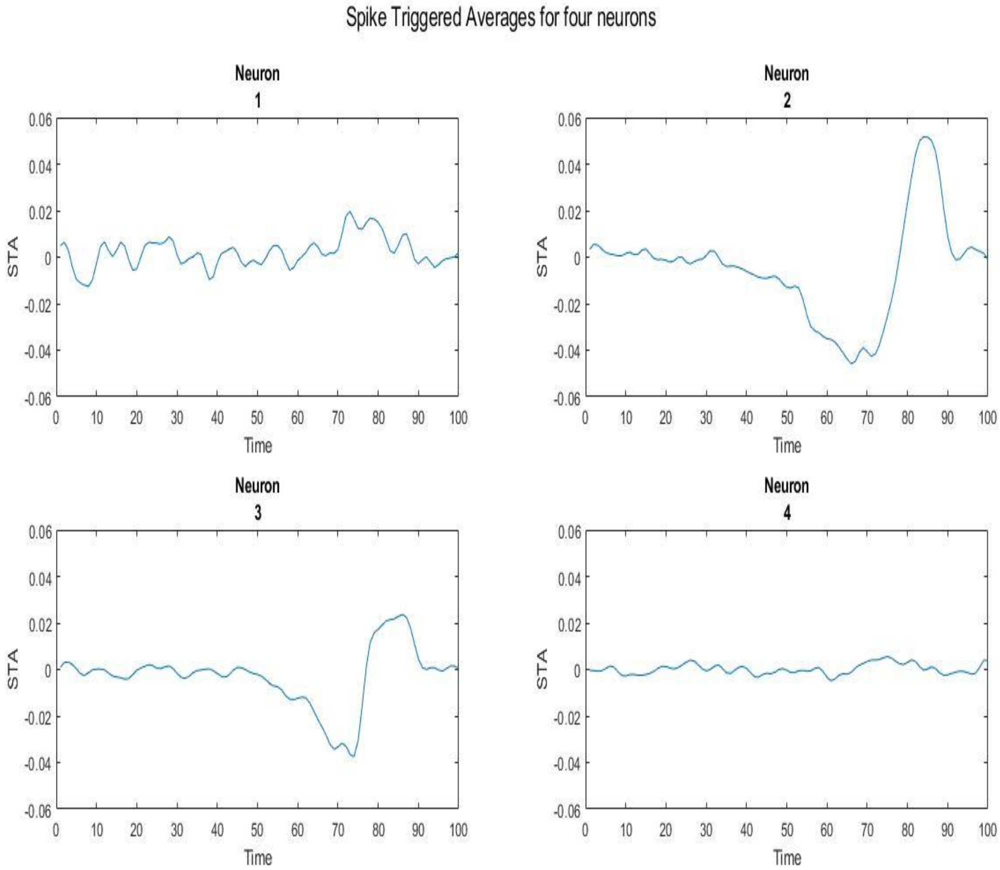

Figure 10: STA for four neurons

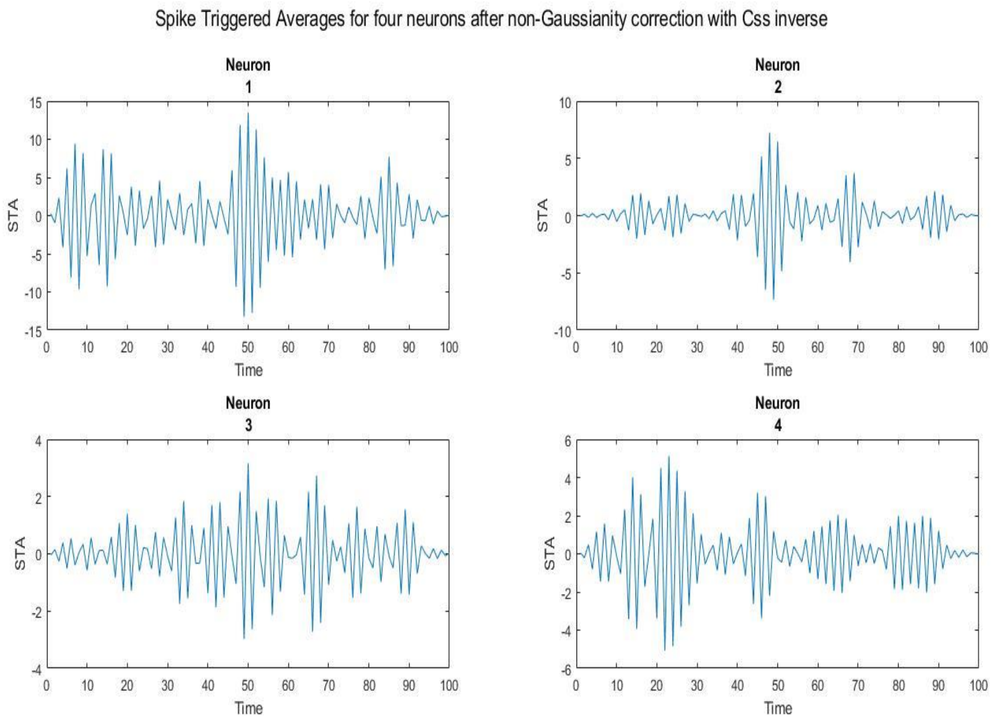

Figure 11: STA for four neurons after CSS inverse correcction

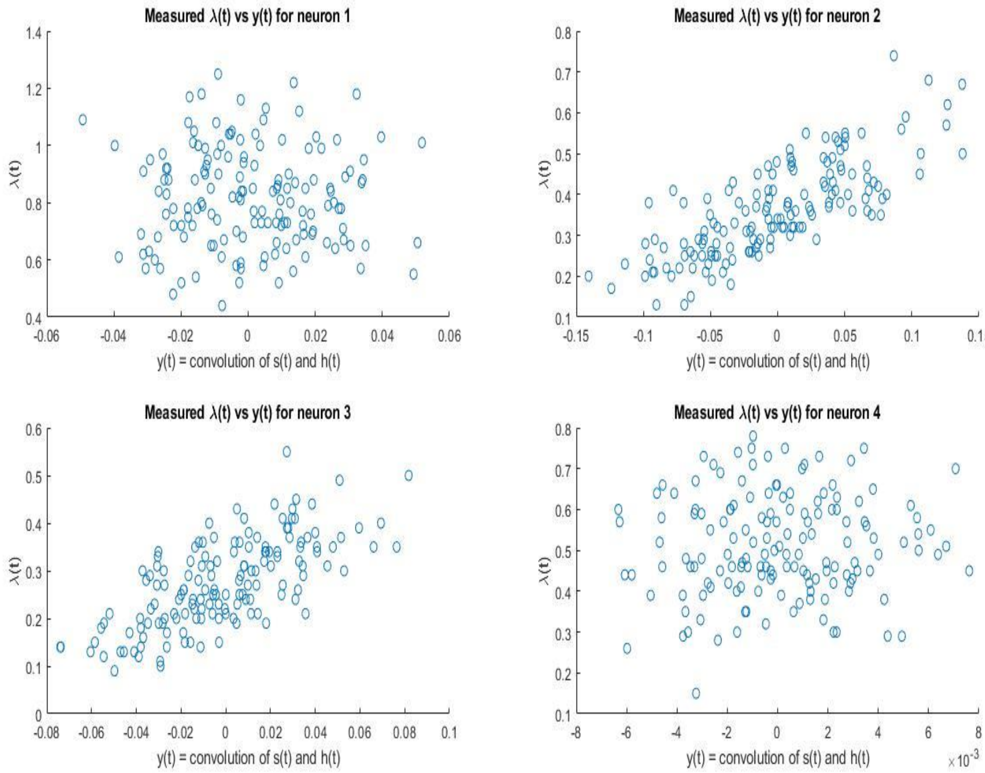

Figure 12: Observed vs predicted values

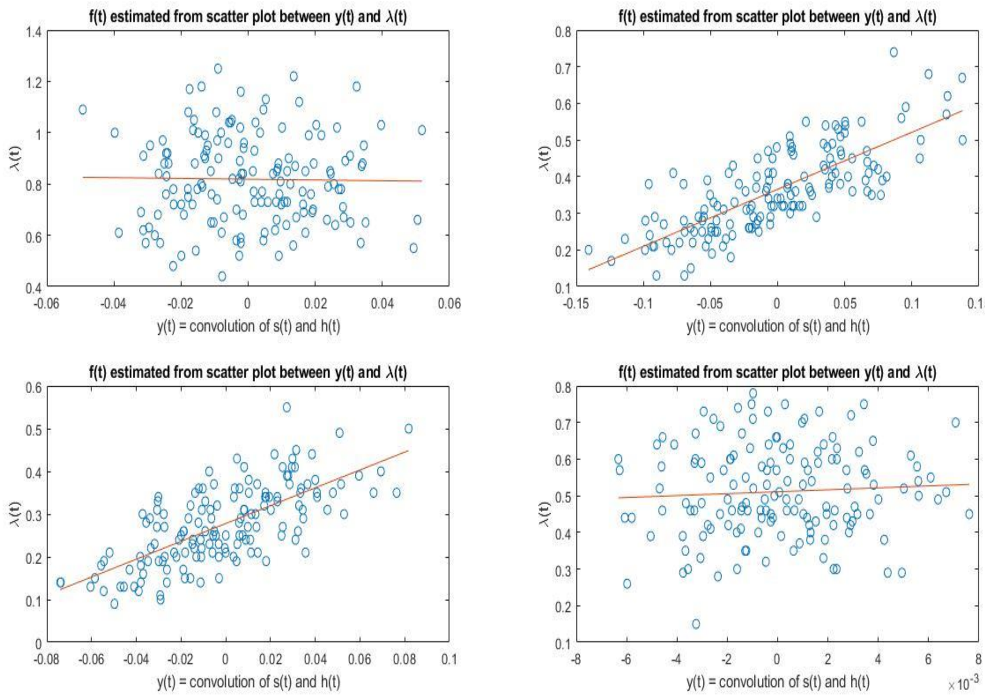

Figure 13: Linear fit on the Fig 12 data

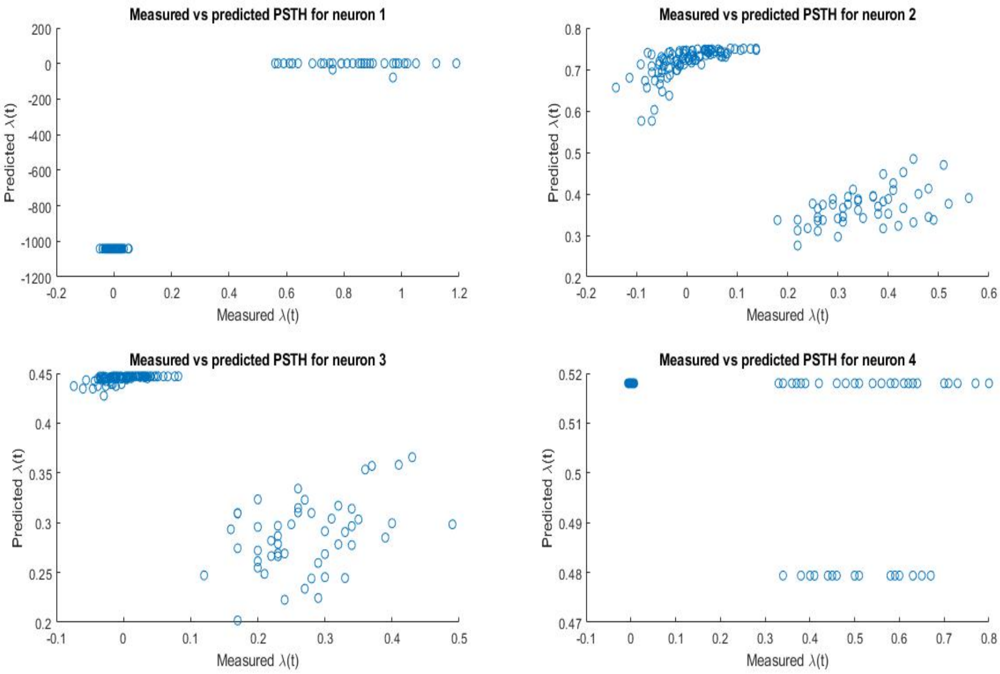

Figure 14: Measured and Predicted Lambda values with the sigmoid fit

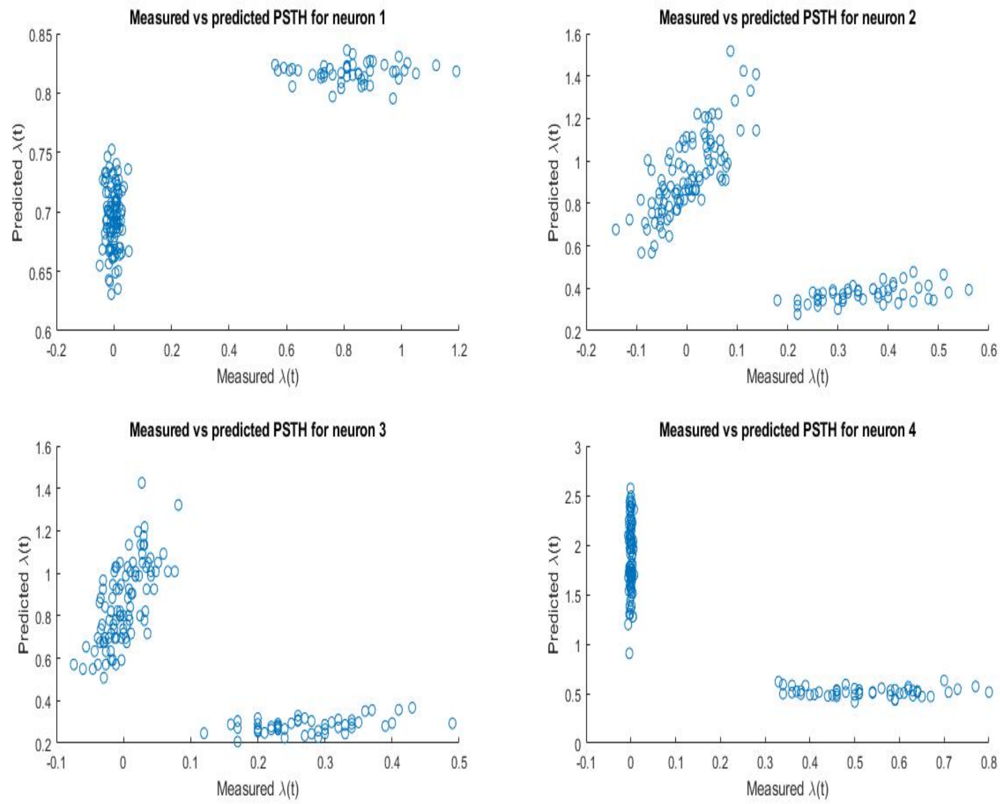

Figure 15: Measured vs Predicted lambda values (PSTH) with linear fit

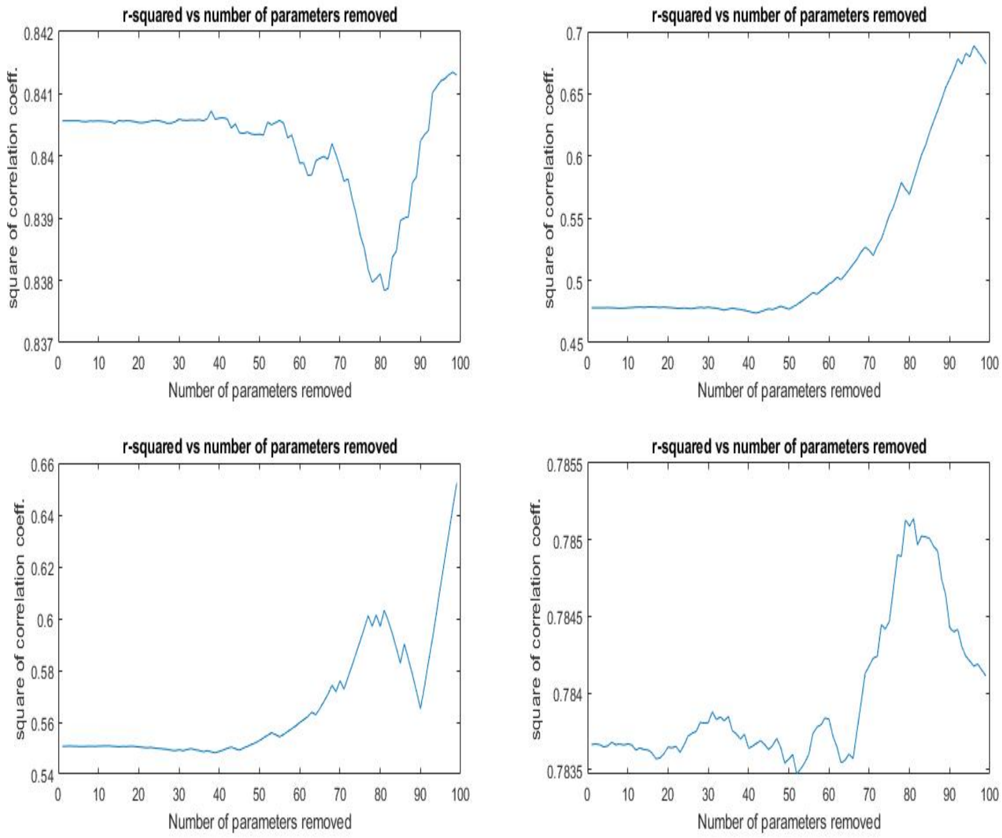

Figure 16: Square of correlation between predicted and observed values vs. the number of points removed

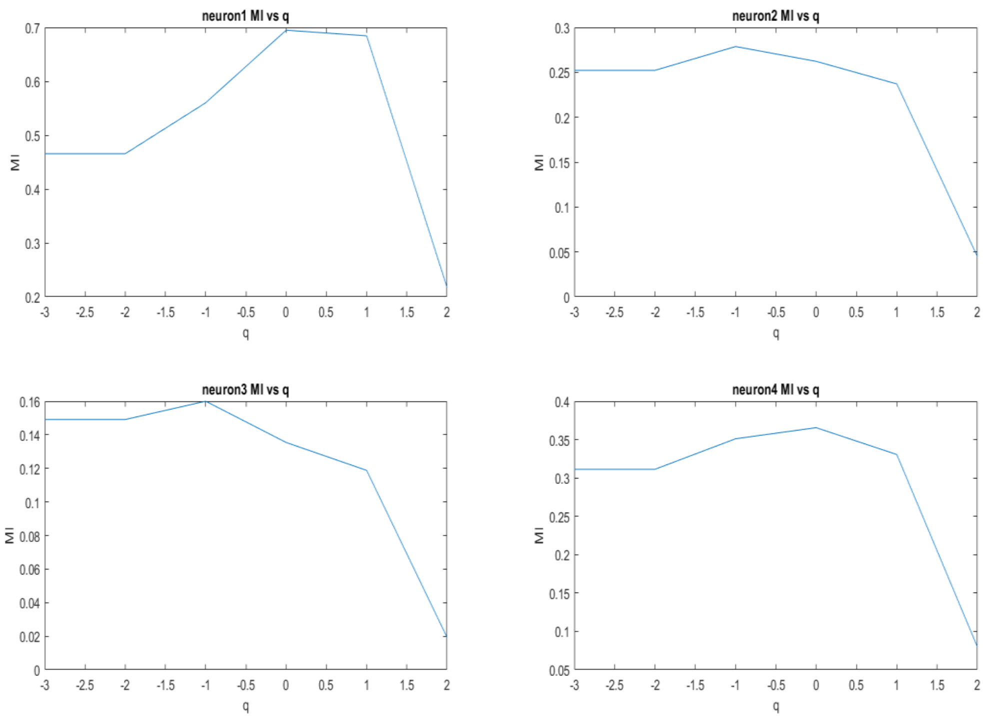

Figure 17: Mutual Information vs. q for all four neurons

Contributions

Sr. No	Name	Roll No.	Contribution
1	Vedic Partap	16CS10053	Q1-6 Answers, Q-7 Code, Report
2	Shounak Sural	16EC10063	Q1-6 Code
3	Prashanthi Silla	16MA20031	Q1-3,5 Code
4	Param Budhraja	17EE30015	Q4,6,7 Answers
5	Ravi Sheoran	14PH20028	Report

Code

The code is available at GitHub. https://github.com/iamshounak/CompNeurosci