J.N. Andrews Honors Program
Andrews University

Proposal for Senior Honors Thesis

HONS 497 Senior Honors Thesis **Credits** __2___ (2 minimum required)

Directions: Please return signed proposal to the Honors Office **at least one week prior to your scheduled meeting with the Honors Council**. This proposal must be accepted by Honors Council the semester before presentation.

Student's Name: Greg Zdor

Primary Advisor: Jay Johnson

Secondary Advisor: none

Thesis Title:

Assessing the Mean Neuronal Firing Rate Information Hypothesis via Mutual Information

Local phone: 269-605-9891 Email: greg@andrews.edu

Expected date of Graduation: December 2018

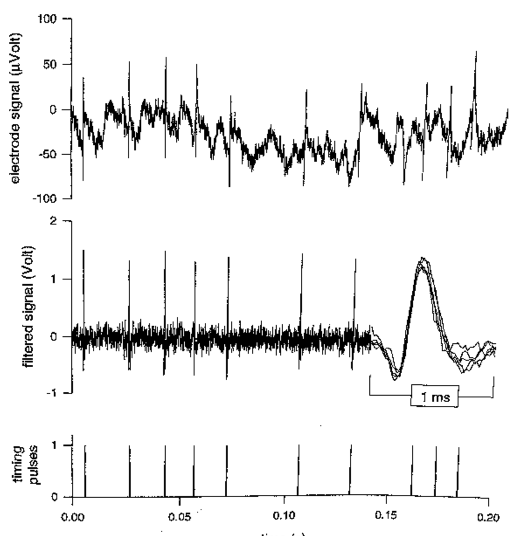
I. Provide goals and brief description of your project or research.

Background and History

Neurons send information from one neuron to the next via electric signals (action potentials). These signals are in the form of slight variations in electric potential, or voltage, which are on the micro volt scale. In 1926 researchers Adrian and Zotterman designed an electrical apparatus for recording the electrical activity in a nerve fiber.[1] However, since the electrical signals from individual cells are very small, these signals needed amplification for analysis, which was provided in the early 1900s by the valve vacuum tube amplifier. Now, with a way to amplify voltage signals, Adrian was able to analysis the signals, the results of his experiments being summarized in his *The Basis of Sensation* (1928). His experiments demonstrated three foundational facts about the neural code.

1. Individual sensory neurons produce action potentials or spikes
2. The rate at which a neuron spikes or fires increases in response to a static stimulus to the neuron
3. If a static stimulus is applied to the neuron for a very long time, the spiking rate decreases, a process known as adaptation

Figure 1: Voltage spikes or action potentials over time (spike trains) for a tungsten wire electrode near a single neuron in the brain of a fly. The top graph shows the raw data, the basic output of neural code. This panel shows the voltage difference between the neuron in the fly's cell, and a ground in the fly's body fluid. The middle panel shows a normalization and filtering of the signal, with the high frequency components separated from the lower frequency spikes (the tallest spikes). The bottom graph shows analogous discriminator circuit generated voltage pulses. [2]



Adrian's experimentation laid the foundation for modern characterization of the neural response. While early experiments measured the response of a neuron by counting the number of spikes in a set period of time upon stimulus application to get a mean rate, today's experiments generally consist of measuring an ensemble of data measurements resulting from an identical stimulus and then performing ensemble averages. If multiple data runs were graphed on top each other, it would be apparent that the spike trains are not identical in each trial, indicating a presence of randomness in the neuron output signal. Looking at the middle

[1] Adrian, E. Zotterman, Y. "The Impulses Produced by Sensory Nerve – Endings: Part II. The Response of A Single Ended – Organ." *The Journal of Physiology*. 61(2). April 23 1926. 151 – 171.
[2] Rieke, Warland, Steveninick, and William Bialek. *Spikes Exploring the Neural Code*. MIT Press 1999. P. 4-5.

graph in Figure 1, it may be seen the spikes do not occur at the exact same spacing in time, evidence of this randomness. Yet as Figure 1 indicates, there is a given time gap at which spikes have a higher probability of firing, though each spike may deviate from that average spacing by some amount. Thus, neural response is related to the average rate at which the spikes fire: the *mean firing rate*. Evidently then the rate at which a neuron fires encodes information about sensory and motor events, which is an accepted fact for neuroscientists since Adrian and colleagues' discovery over 90 years ago. More recently, however, as Stein and colleagues point out, "others have questioned whether rate is the only parameter that encodes information about sensor and motor events."[3] This query leads to the topic of my research.

Goal and Brief Description

My Honors Thesis involves evaluating the information content of a data set of single marmoset neuron's spike trains in response to an applied auditory stimulus. As noted prior, spike train neural encoding information is well known to derive from the mean firing rate. My goal is to assess the null hypothesis that *all information content of the considered data set comes solely from the mean firing rate.* Briefly put, my project consists of two parts: evaluating the actual data set and then comparing this with results from surrogate data.

The first stage of my project will be to find the total information content of the *actual* data set. I will do this by first finding the relative spacing of the individual spikes within the data set, from which I will be able to calculate the probability density functions for each spacing. I will then calculate the entropy[4] for the ensemble of inter – space gaps (ISG), from which the mutual information (MI)[5] may be calculated for that given ISG. Summing MI over the complete data set gives the total information present. This concludes the first part.

The second part of my project is to perform surrogate data testing to provide comparison data. I will create surrogate data via one of two ways: (1) either assume a given mean firing rate and randomly generate data via a Poisson distribution using the average ISG width, or (2) take the complete array of ISG data and randomly shuffle it. Upon creating comparison data, I will then repeat the above steps in part one, except assuming the initial null hypothesis. This second step I repeat 100 times, from which I will get an average mutual information under the mean firing rate assumption. The final step is comparing whether the total MI from the actual data set is larger or smaller than the averaged total MI from the surrogate data set, which will provide a proper way to assess my initial null hypothesis.

II. Outline your methodology. Please be specific. How does this achieve your goals and how reliable is it?

Methodology:

I will be working in the computer program MATLAB, a numerical computing environment, for the analysis of data. The data is in the form of .wav and .mat files in MATLAB which I have from Dr. Johnson. Dr. Johnson procured these data sets from Dr. Ross Snider of Montana State University's Electrical Engineering Department. Dr. Snider acquired this data from John Hopkins University a number of years ago as part of a federally funded study. Federal funding providing the origins of this data ensures its acquirement from the marmoset was done in accordance to the Animal Welfare Act. I will perform research on the Physics Department's research computers in Johnson's computer research lab in Haughey Hall.

[3] Stein, R. Gossen, R. Jones, K. "Neuronal Variability Noise or Part of the Signal?" *Nature Reviews: Neuroscience.* Vol 6. May 2005. P. 389.

[4] Entropy defined as the average amount of information conveyed by an event, when considering all possible outcomes. The more uncertain the event, the higher the entropy.

[5] A working definition of MI is a measure of how much one given variable tells us about another one.

The first step of my project is to determine what constitutes a "spike" and what is background noise. I will do this by either (1) identifying the peak voltage and assigning it to "spike" or (2) using the repeated waveform of the spikes as the characterizing trait by which to define them. Upon defining spikes, I will then determine the inter – spike gap (ISG) between each spike. This will require defining a certain spike width, whereupon all other data in between spikes is defined as ISG data. Between each spike is an ISG, so if I have N number of spikes, I will have N – 1 number of ISGs. Figure 2. shows a representative ISG denoted by delta: Δ, while the # denotes a spike. In MATLAB, I will assign these ISGs to an array N-1 wide.

Step two is where I define a probability distribution function for each ISG. To do this, I graph the first ISG Δ_1 versus Δ_2, Δ_2 (on x – axis) versus Δ_3 (y – axis), and so on for the complete data set, graphing the current versus the next ISG value. This results in a 2 – dimensional graph of scattered data points. Data binning is then done to the data set. This can be best imagined as drawing a crisscrossing gridline atop the graph where all the data points lie in one of the square bins created by the imaginary overlain grid. This process of binning the data is done in MATLAB. In order to bin the data, I will first have to select the bin size, which will be dependent on the amount of data. With each data point in a bin, I then will assess the probability distribution function of each of the bins in each dimension.

Figure 2: Below is a enlargement of Figure 1, with representative inter – spike gap (ISG) marked by Δ_1, Δ_2, and Δ_3 while representative spikes are marked with #1, #2, and #3.

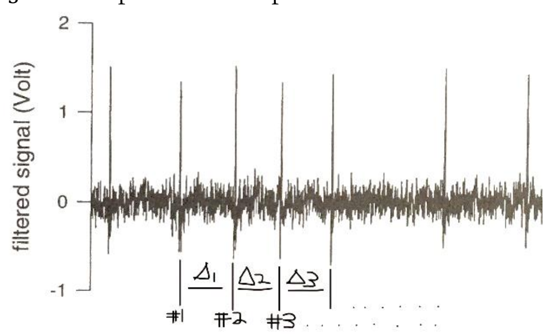
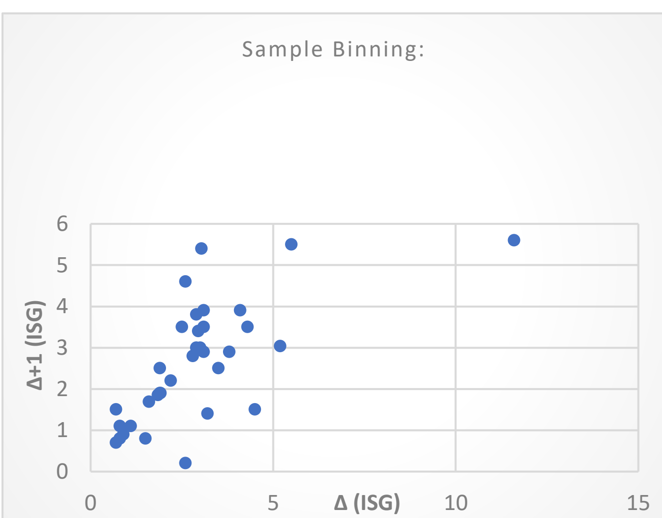


Figure 3: This figure shows a sample of arbitrary data that has been binned.



Each data point in a graph like Figure 3 has both x (horizontal) and y (vertical) components. This is important when considering the probability, since it if the total probability in a given direction is considered, then there is two dimensions or variables to which to assign probability, X for the x dimension and Y for the y dimension. Each of these variables is comprised of the individual probabilities of that given variable's probability in its respective bin summed. Equations 1 and 2 below represent the *individual* probability of a given bin in its respective direction.

$$\sum_{x(1)}^{x(M)} \frac{\hat{x}}{M} = P_{bin}(\widehat{x}) \qquad \text{Equation 1}$$

$$\sum_{y(1)}^{y(N)} \frac{\hat{y}}{N} = P_{bin}(\widehat{y}) \qquad \text{Equation 2}$$

Now in order to get the information content of the data, mutual information (MI) from information theory is used. MI may be calculated from the entropy of the given variable of consideration, X and Y in this case.[6] Denoted by H, the definitions of entropy in terms of the probability distributions follow:

$$H(X) = -\sum_{m=1}^{m=M} P_{(bin\,m)}(\widehat{x}) \log(P_{(bin\,m)}(\widehat{x})) \qquad \text{Equation 3}$$

$$H(Y) = -\sum_{n=1}^{n=N} P_{(bin\,n)}(\widehat{y}) \log(P_{(bin\,n)}(\widehat{y})) \qquad \text{Equation 4}$$

$$H(X,Y) = -\sum_{n=m=1}^{n=N,m=M} P_{(bin\,n,m)}((\widehat{x)(y})) \log(P_{(bin\,n,m)}((\widehat{x)(y})) \qquad \text{Equation 5}$$

Mutual information – MI – may now be calculated:

$$Mutual\ Information = MI = H(X) + H(Y) - H(X,Y) \qquad \text{Equation 6}$$

The calculations of equations 1-6 are calculated in MATLAB. The final step consists of creating surrogate data and comparing this with the results of first calculation of MI.

Such comparison data may be created several ways, including the following: (1) either assume a given mean firing rate and randomly generate data via a Poisson distribution using the average ISG width, or (2) take the complete array of ISG data and randomly shuffle it. Upon creating comparison data, the above steps to calculate MI are repeated, except now assuming the initial null hypothesis. I repeat creating surrogate 100 times from the original data set and calculate MI a corresponding number of times, from which I get an average MI value for the surrogate data. This I then compare with my original calculated value of MI from where I can properly assess whether my initial null hypothesis.

III. Explain in what sense your project is original, unique, or beyond normal senior expectations. How does it relate to current knowledge in the discipline?

As an Engineering major with a concentration in Electrical and Computer Engineering and with Mathematics and Physics minors, I am required to do a senior capstone Senior Engineering Design Project. The project proposed here is wholly separate from what I will be doing for my Senior Engineering Design Project, rendering it beyond normal senior expectations. This project finds relevance to me since my major and minors cover physics, mathematics, and signal analysis.

---

[6] Wing, S. Johnson, R. Camporeale, E. Reeves, G. "Information Theoretical Approach to Discovering Solar Wind Drivers of the Outer Radiation Belt." *Journal of Geophysical Research: Space Physics*. 10.1002/2016/JA022711. P. 4.

A literature review of Andrews University James White Library's databases along with Notre Dame's Hesburgh library did not show information theory and mutual information applied to a single neuron's information content in contrast with the MI due surrogate data from mean firing rate. On a larger scale this project finds relevance in the larger community of neuroscience, as the researchers are asking "whether rate is the only parameter that encodes information about sensory and motor events."[7] Information content of neurons is of interest, as neuroscience is still seeking to characterize the neural response fully.[8]

IV. Include a substantive annotated bibliography of similar or related work.

Wing, S. Johnson, R. Camporeale, E. Reeves, G. "Information Theoretical Approach to Discovering Solar Wind Drivers of the Outer Radiation Belt." *Journal of Geophysical Research: Space Physics*. 10.1002/2016/JA022711. P. 4.

Simon Wing, Jay Johnson, Erico Camporeale, and Geoffrey Reeves provide a useful research article on how they used mutual information, conditional mutual information, and transfer entropy to analyze the nonlinear dependence relationship of solar wind velocity of the magnetosphere on geosynchronous MeV electron flux peaks. For the most part this article dealt with magnetosphere science, naturally, a topic unrelated to my research. However, early on authors Wing et al provide an excellent summary of how they applied information theory to radiation belt MeV electron data. This article proves useful to me as it provides an excellent basis from which to cite the equations of MI, TE, and CMI.

Stein, R. Gossen, R. Jones, K. "Neuronal Variability Noise or Part of the Signal?" *Nature Reviews: Neuroscience*. Vol 6. May 2005. P. 390-7.

Richard Stein, Roderich Gossen, and Kelvin Jones present a review of neuronal variability, asking the question whether noise in neuron spike trains holds important information or is merely noise. I found this as a pertinent article as it provided a review of a number of the views on meaning firing rate and information content. The authors put forth that both temporal and rate coding (mean firing rate) are present to different degrees in the nervous system. Particularly interesting I found the fact that changes in signal frequency correlate with increased information content, even if the frequency decreases, while constant signal frequency is proportional to a decrease in information content. Overall, this article provided a helpful background to the issue of information content in neural networks.

Borst, A. Juergen, H. "Effects of Mean Firing on Neural Information Rate." *Journal of Computational Neuroscience*. 10, 213-221, 2001.

Authors Juergen and Borst investigated the effect of the mean firing rate on the information rate of motion sensitive neurons in flys. Key to their work was the deployment of entropy to determine information content. The results indicated that information rates increased in all their test conditions with increase in mean firing rate. To me, this article pointed strongly to the historical idea that the mean firing rate holds the predominant amount of information about neural encoding. However, finer nuances lost in averaging and in ignoring the order of spikes I see as points this article does not address. I found this article relevant to my research as it

[7] Stein, R. Gossen, R. Jones, K. "Neuronal Variability Noise or Part of the Signal?" *Nature Reviews: Neuroscience*. Vol 6. May 2005. P. 389.

[8] Stein, R. Gossen, R. Jones, K. "Neuronal Variability Noise or Part of the Signal?" *Nature Reviews: Neuroscience*. Vol 6. May 2005. P. 390.

provided a basis outside of *Spikes* of relatively current data pointing heavily towards meaning firing rate information content.

Lundstrom, B. Famulare, M. Sorensen, L. Spain, W. Fairhall, A. "Sensitivity of Firing Rate to Input Fluctuations Depends on Time Scale Separation Between Fast and Slow Variables in Single Neurons." *Computational Neuroscience*. Vol. 27: 277 – 290. 2009.

Fairhall and colleagues separated neurons into three general categories in accordance to how their mean firing rate varied on input mean and variance. They then found three biochemical explanations for how a neuron could possibly increase it's sensitivity to input fluctuations – one including a change in conductance. This article provided interesting insights broadening my exposure to the biological side of computational neuroscience.

Rieke, Warland, Steveninick, and William Bialek. *Spikes Exploring the Neural Code*. MIT Press 1999.

This 395-page tome starts with the basics of Adrian's perfunctory experiments in 1926 and lays a rigorous mathematical foundation for today's current understanding of the neural code in neuroscience. I found this book the most helpful resource I have had in this project as it starts on a basic level, elucidating how the mathematics behind the current models came to be and why. Moreover, excellent diagrams, illustrates, and explanations added to make the book a great read for not only someone doing computational neuroscience research but for even the common reader.

V. Provide a statement of progress to date and list the research methods coursework completed.

This project has not been started.

Research methods coursework:

Linear System Analysis (signal analysis in MATLAB)
Theoretical Mechanics (MATLAB coding)
Probability Theory with Statistical Application
Electronics I-II, Circuit Analysis***
Physics for Scientists and Engineers I-II***
Logic Circuit Design***

***these classes provided background to my understanding this research as a whole

Department Chair Approval

- **This student's performance in his/her major field is acceptable.**
- **He/she has completed the requisite research methods coursework for the research to be pursued.**
- **I understand that he/she plans to graduate with Honors.**

____

Department Chair (signature)

Research Advisor Approval

I have read and support this proposal: ______________________________________

Primary Advisor (signature)

I have read and support this proposal:

____

Secondary Advisor (signature)

If human subjects or if live vertebrate animals are involved, evidence of approval from the Institutional Review Board or an Animal Use Committee is needed through the campus scholarly research offices (Ext. 6361).