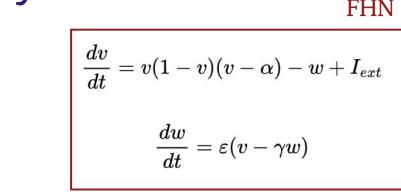
# Dexter Theisen - Hunter College High School

# Introduction

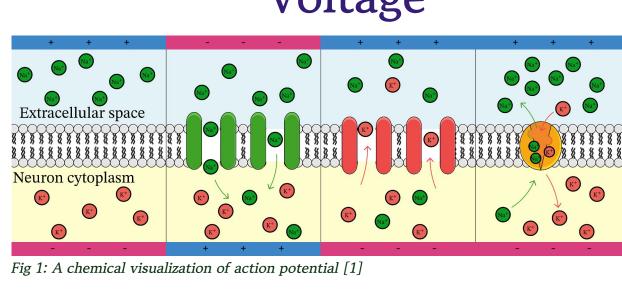
Accurately modeling the action potential of brain cells is critical to advancing neuroscience by developing computational methods to analyze cerebral activity. This project investigates the estimation of parameters in the FitzHugh-Nagumo (FHN) system, a simplified yet robust model used for excitable systems such as neurons.

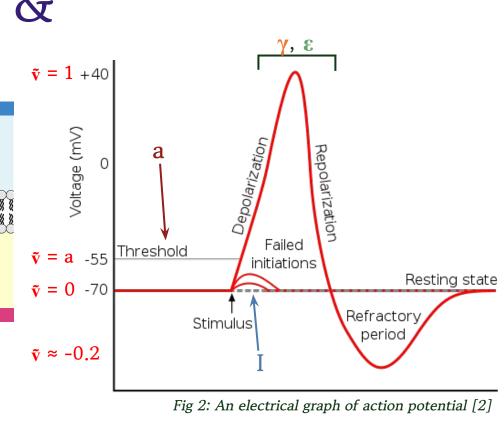
# FitzHugh-Nagumo System

- 1962: Richard FitzHugh and Jinichi Nagumo created a circuit of two differential equations used to model an excitable system (i.e. a neuron)
- They derived their equations from the more general Van der Pol Oscillator (1927)
- v represents the voltage variable, and w represents a blocking mechanism (how open or closed chemically pathways across the neural membrane are)



# Action Potential & Voltage





1. a represents the voltage threshold at

2. Gamma ( $\gamma$ ) controls how quickly w

Epsilon (ε) modifies how quickly w

which a neuron fires

responds to changes in v

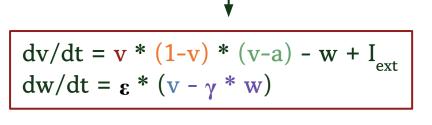
4. I (shorthand I) represents an

external stimulus (i.e. pain)

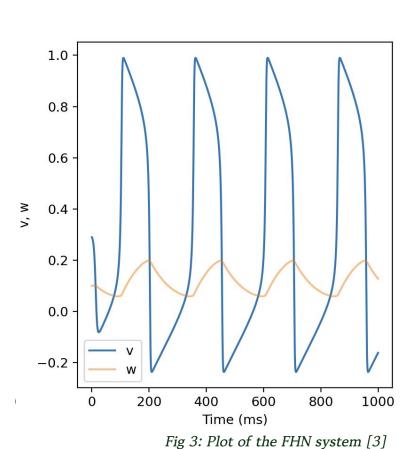
follows v

# Mathematical Patterns

- 1. More voltage → faster increase in voltage 2. There is a max voltage, namely at v = 1
- 3. If v<a: neuron does not fire
- 4. If v>a: neuron does fire; positive feedback
- 5. The blocking mechanism increases in strength as v increases (in order to bring v down after it peaks)
- 6. The blocking mechanism can only get so strong



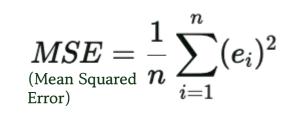
# Objective

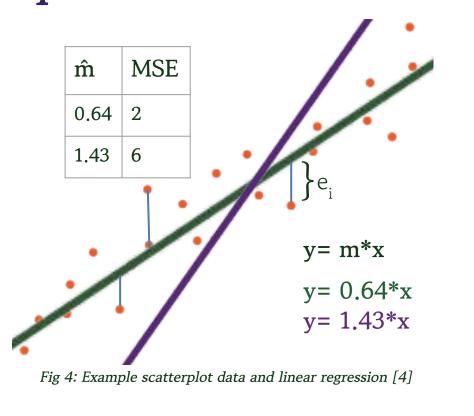


- To ingest data and determine the parameters (a,  $\gamma$ ,  $\epsilon$ , and I) that match the
- Use the estimated parameters to integrate the FHN system and predict the future state of the neuron

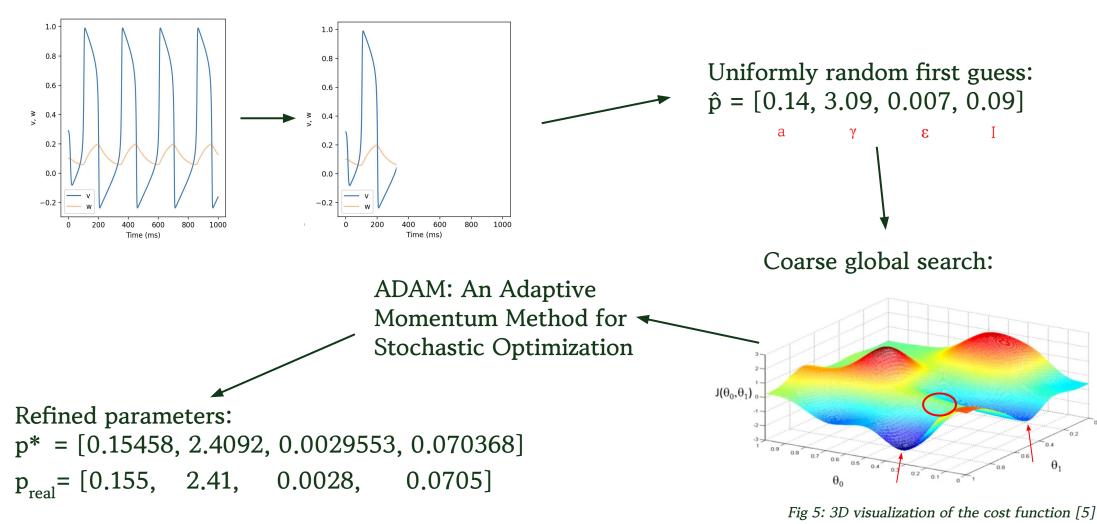
# Regression: Mean Squared Error

- An error function is used to numerically evaluate how "incorrect" a certain guess for parameters is
- The goal of regression is to minimize the error function



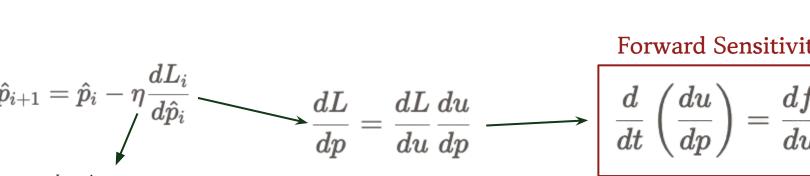


# Regression Flowchart for FHN



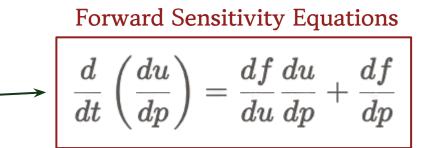
# Coarse Global Search

- Objective: determine how to update parameter guess (p vector) based on the current guess
- Method: determine how the loss changes with respect to each parameter, and update each parameter in the right direction

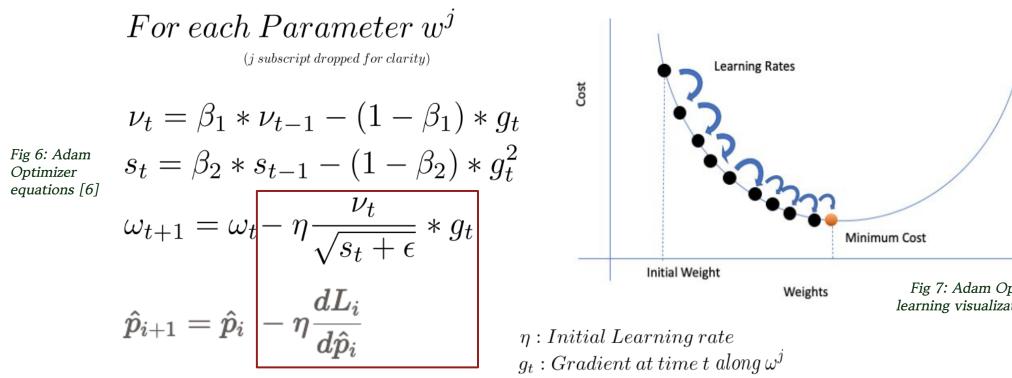


# Definitions:

Definitions: 
$$ec{f} = egin{bmatrix} v(1-v)(v-lpha) - w + I_{ext}, \ arepsilon(v-\gamma w) \end{bmatrix}$$
  $ec{u} = egin{bmatrix} v(v-\gamma w) & ec{p} = [lpha, \gamma, arepsilon, I] \end{bmatrix}$ 

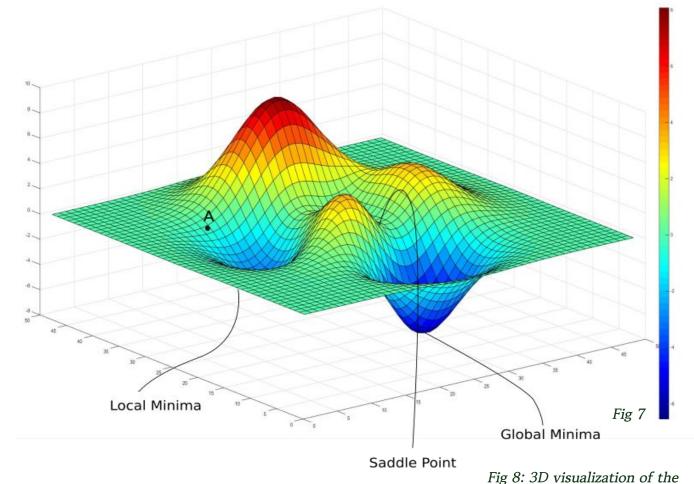


# Adam Optimizer



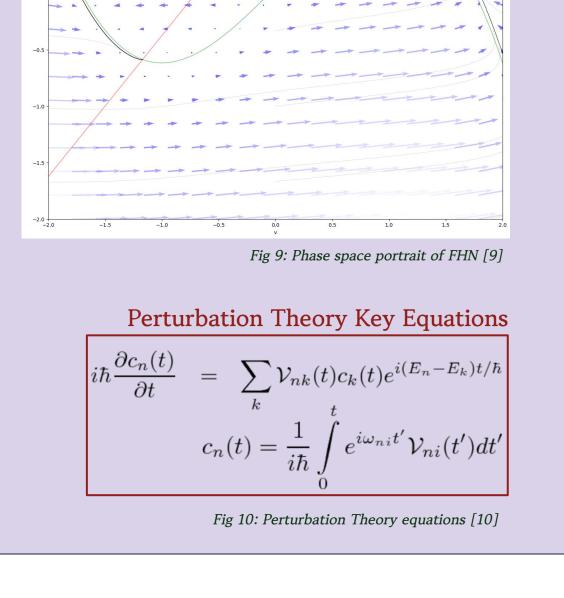
# Why Use Adam?

• Because the Adam Optimizer is based on the physics of a rolling ball, it maintains a type of momentum while rolling down the "cost hill," allowing the "ball" (our best guess) to avoid local minima

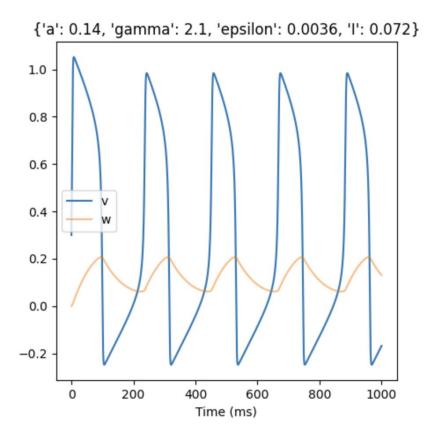


# Boundaries

- In nature, neurons have a resting electrical energy of -70mV and a peak energy of 40mV
- During hyperpolarization the energy levels drop to around -90mV • V is scaled such that:  $-70 \text{mV} \Rightarrow \text{v} = 0$  and  $40 \text{mV} \Rightarrow \text{v} = 1$
- Therefore, generally:  $v \in [-0.2, 1]$
- w represents the openness of chemical pathways
- It is hard to conceptualize a negative value for w
- (as 0 already implies that the channels are fully closed)
- Thus, a lower bound is:  $w \ge 0$
- a is the threshold for neural firing
- Biologically, this value lies within in the range -50mV and -55mV • Scaling to the appropriate range yields: ~0.13 < a < ~0.18
- To ensure stability in the FHN, the cubic nullcline for dv/dt and the
- linear nullcline for dw/dt must intersect • By letting  $G(v) = F(v) - v/\gamma$  and plugging in G(0) and G(1), we conclude
- (assuming an intersection point, and thus G(0)>0 and G(1)<0):  $0 < I < 1/\gamma$ • It turns out that  $1/\gamma = m$  (the maximum slope of the cubic nullcline that still
- ensures an intersection)  $\approx 0.28$
- Thus:  $0 < I < \sim 0.28$ • And similarly,  $0 < \gamma < \sim 3.57$
- (The lower bound for  $\gamma$  is 0, as a negative  $\gamma$  would cause w –
- and subsequently v to diverge)
- Following Perturbation Theory, in the original FHN equations,
- v is the "fast" variable and w is the "slow" variable • Therefore the timescale of v must be greater than the timescale of w
- The characteristic timescale  $\tau_{v}$  can be estimated as  $\sim 1/\max(v(1-v)(v-a)) \approx 8.3$
- The characteristic timescale  $\tau_{w}$  is set by  $\varepsilon^*$  (v  $\gamma^*$  w), but because its order of
- magnitude is typically O(1), we conclude  $\tau_{...} \sim 1/\epsilon$ • Given  $\sim 8.3 > 1/\epsilon$ , we conclude:  $\epsilon < \sim 0.12$



#### Test 1

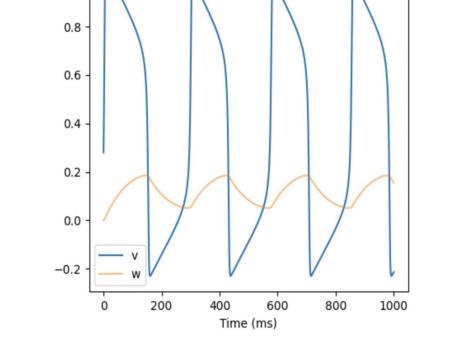


REAL:

# {'a': 0.1695, 'gamma': 2.8, 'epsilon': 0.002, 'I': 0.095}

200

Test 2



Test 3

{'a': 0.15, 'gamma': 3.11, 'epsilon': 0.0026, 'l': 0.0615}

# Data Tables

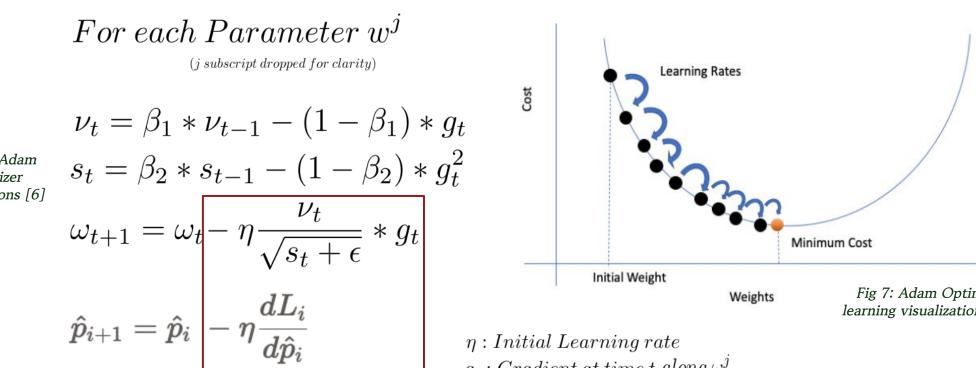
400

600

a	γ	3	I
0.140038	2.10137	0.00359978	0.0719515
0.140048	2.10131	0.00359978	0.0719561
0.140031	2.10141	0.00359978	0.0719483
0.140047	2.10131	0.00359978	0.0719558
0.140024	2.10145	0.00359979	0.0719555
0.14037	2.10138	0.00359978	0.0719511
0.140039	2.10137	0.00359978	0.0719520
0.14	2.1	0.0036	0.072
	0.140038 0.140048 0.140031 0.140047 0.140024 0.14037 0.140039	0.140038 2.10137   0.140048 2.10131   0.140031 2.10141   0.140047 2.10131   0.140024 2.10145   0.14037 2.10138   0.140039 2.10137	0.140038 2.10137 0.00359978   0.140048 2.10131 0.00359978   0.140031 2.10141 0.00359978   0.140047 2.10131 0.00359978   0.140024 2.10145 0.00359979   0.14037 2.10138 0.00359978   0.140039 2.10137 0.00359978

#### 01 0.09502 12 0.09510 210 0.09508 07 0.09507 0.09507 06 0.09498 99 0.09500 0.095

# T1: | 0.150012 | 3.110114 | 0.0026097 | 0.06143 T2: | 0.150100 | 3.110324 | 0.0025968 | 0.061493 T3: 0.150085 3.110098 0.0026043 0.061502 T6: 0.149941 | 3.109893 | 0.0025899 | 0.061520 T7: 0.150111 3.110203 0.0025913 0.06108 | REAL: | 0.15 | 3.11 | 0.0026 | 0.0615



# $\beta_1, \beta_2: Hyperparameters$

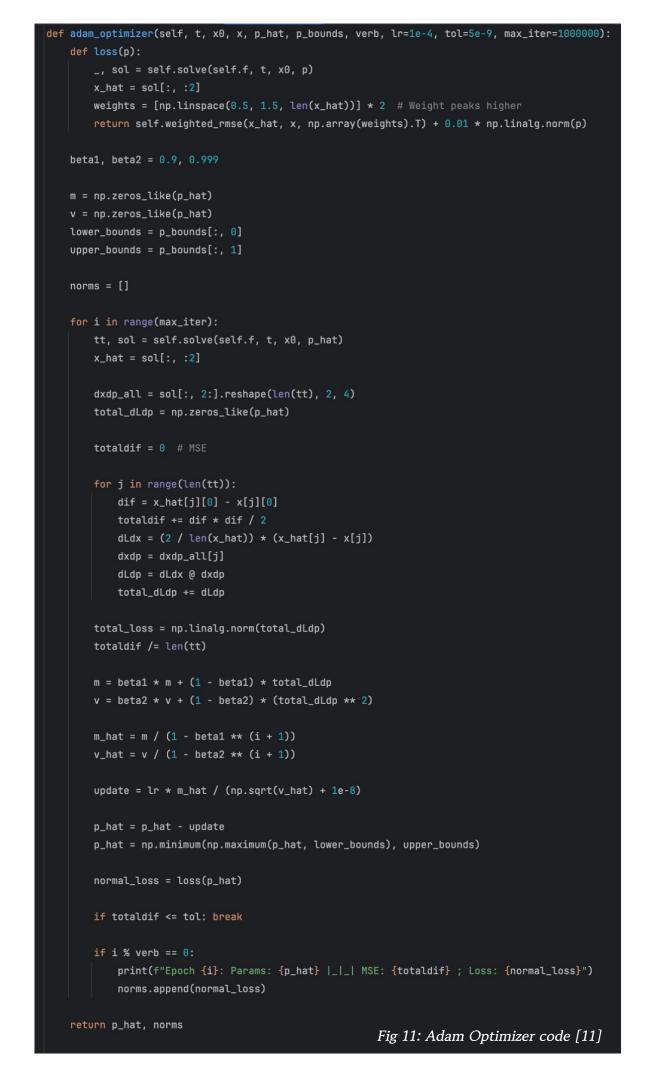
• Downside: this process often takes longer, as more operations are calculated and more confidence is required before the program confidently terminates

 $\nu_t$ : Exponential Average of gradients along  $\omega_i$ 

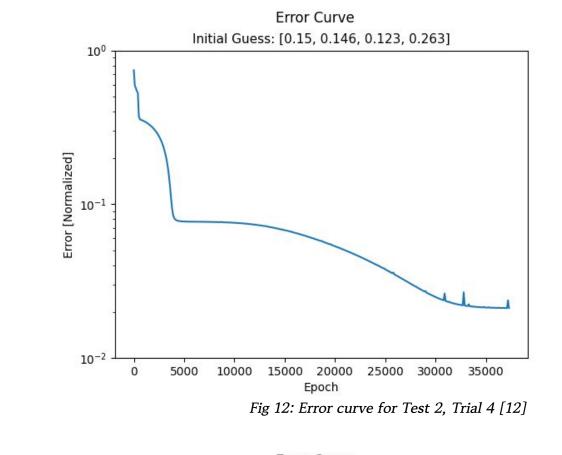
 $s_t: Exponential \ Average \ of \ squares \ of \ gradients \ along \ \omega_j$ 

• Hence, this project deployed a combination of the normal Stochastic Gradient Descent (a first, coarse global search) and Adam Optimizer (a second, fine tuned search), to achieve the best of both algorithms

# Code Sample



# Sample Error Curves



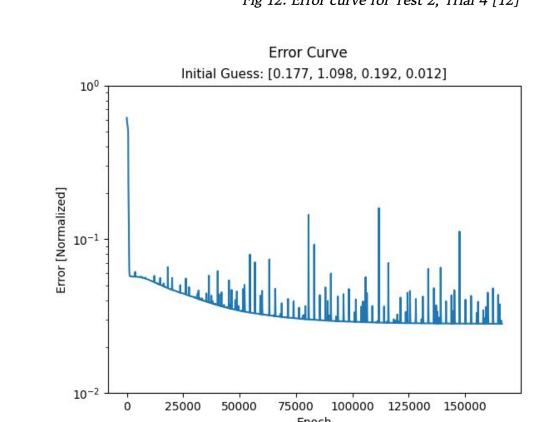


Fig 13: Error curve for Test 1, Trial 6 [13]

#### Conclusion

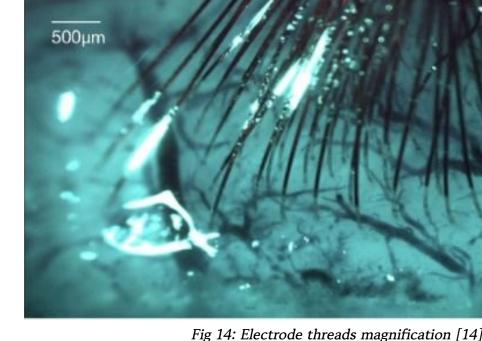
This project has demonstrated that it is possible to fit the FHN model to data with unknown parameters. By generating synthetic data and employing a combination of constrained optimization and a modified gradient descent algorithm, the parameters were accurately recovered within biologically plausible bounds. The results consistently achieved a mean squared error below 10<sup>-9</sup>, which is a high enough accuracy to predict a neuron's state tens of seconds into the future.

# Forward-looking Applications

- Electrotherapy without a reliance on human test subjects
- a. By modifying the I<sub>ext</sub> variable, we can test the neural reaction to given stimuli, ensuring safety (ie verifying the graph of neurons remains stable) before treatment is
- 2. Prosthetics

3. Neural Modeling

a. By using this work to predict the future electrical current of a given neuron/neural network, we can improve upon the ability of prosthetics to act with as little latency as possible (in order to better mimic human body parts)



- a. By programmatically linking multiple neurons together and using the same math/ideas proposed by this project, we can hopefully model human brains on the neuron level (using Neuralink electrode threads), as opposed to using clumps of neurons for EEG data

# Future Work

- 1. Estimate the parameters for each neuron in a larger neural network
  - a. This, in combination with Neuralink's continued refinements in measuring neurons more precisely, will hopefully in the future allow for low-level accurate neural mapping as discussed above
- 2. Attempt to estimate the parameters for real-world data
- a. This will test the viability of using the FHN system for modeling human neurons
- 3. Omit the w variable as input data and only use v data, as this is more likely represents the availability of data in real-world situations
- a. This requires additional math such as Inhomogeneous-Poisson processes and spike trains, but seems feasible

#### Acknowledgments

- I would like to express my deepest gratitude to the Hunter College High School Math and Science Departments for the exemplary instruction and encouragement they have bestowed on me to pursue my passions for math, science, and coding. I would particularly wish to thank Mr. Cocoros (Calculus BC teacher), Mr. Owens (Mathematics Department Chair), Mr. Park and Ms. Gonzalez (Advanced Science Research Seminar Chairs) for helping me reach this milestone in my research
- Also, a big thank you to my math lab director and friend, Andrew Sabatino, for instilling in me the love of STEM at a young age and teaching me how to continuously nurture it
- Finally, I am immensely thankful for the thoughtful counsel of my project advisor, Erick Ruiz. I could not have found my way through the dozens of hours of code and math equations to complete this project without his patient and sage guidance

#### Sources

#### General:

- . All plots in the same style as Figure 3 were generated by the finalist using Matplotlib
- 2. All data tables were generated by the finalist

#### Specific Graphic Reference List:

- 1. Nawga. The Nerve Impulse. https://www.nagwa.com/en/explainers/494102341945/
- 2. Khanna, Aarushi. Action Potential. https://teachmephysiology.com/nervous-system/synapses/action-potential/
- 3. Image created by the finalist using Matplotlib
- 4. Image created by the finalist using Desmos Graphing Calculator
- 5. Manikandan, Bala. Machine Learning: Gradient Descent. https://blogtech.online/machine-learning-gradient-descent-1/
- 6. Kingma, Diederik and Ba, Jimmy. Adam: A Method for Stochastic Optimization. https://arxiv.org/abs/1412.6980
- 7. Ghosh, Bhaskar. Adam Optimizer. https://www.researchgate.net/figure/Stochastic-Gradient-Descent\_fig4\_344544069
- 8. Kathuria, Ayoosh. Intro to optimization in deep learning: Gradient Descent. https://www.digitalocean.com/community/tutorials/intro-to-optimization-in-deep-learnin g-gradient-descent 9. Nebula, Cosmia. Fitzhugh-Nagumo Model.
- https://en.wikipedia.org/wiki/FitzHugh%E2%80%93Nagumo\_model#/media/File:FitzHugh -Nagumo\_Model.png 10. Branson, Jim. Time Dependent Perturbation Theory.
- https://quantummechanics.ucsd.edu/ph130a/130\_notes/node38.htm 11. Image (and the code contained within it was) created by the finalist
- 12. Image created by the finalist using Matplotlib
- 13. Image created by the finalist using Matplotlib
- 14. Servick, Kelly. Brain Computer-Interface.
- https://www.science.org/content/article/elon-musk-s-startup-eyes-human-testing-brain-c omputer-interface