A novel causal modeling algorithm uncovers predictive coding properties of *c. elegans*

<u>Background</u>
Predictive coding is a theory in theoretical and cognitive neuroscience which aims to prove that the core function of the brain is to minimize prediction error. The central idea is that the brain is organized in a hierarchy of layers. Each layer makes a prediction about the activity of the layer below it. The downward predictions are compared to the activity of the layer, the difference is called prediction error (Friston, 2003). In other words, at each layer there is activation information that is received and a pre-existing expectation of what is to be received. The difference, prediction errors, at each level is what is fed to the level above it. As a result, there is an efficient hierarchy of perceptual errors which gets passed up. These predictions happen on a variety of scales ranging from fine grain sensory level details such as simple shapes or features of objects, to whole scenes and whole objects, to higher level thinking. The goal of the system is to minimize the difference between actual activity and predicted activity at each level, in other words minimize prediction error, the act of doing this minimization results in inference and learning (Millidge, 2021). This is accomplished by performing a procedure called variational inference.

Predictive coding can help explain how a neuron can respond to a stimulus before the stimulus even occurs (Palmer, 2014). In other words, **an organism can learn to anticipate sensory stimuli in its environment that it has not yet experienced** but learned from previous instances to expect. A single neuron and therefore an entire organism can **make predictions about the future values of sensory inputs**.

For example, think of a baseball player swinging his bat at a pitch. The time it takes for his retina to process the flying ball will be too long for him to even perceive the ball's speed and swing his bat accordingly. His retina predicts the image, before the image is even processed!

The existence of predictive coding in this capacity has been investigated predominantly in complex mammalian systems, with the first proposed model explaining perceptual phenomena in the human retina (Srinivasan, 1982). In recent years there has been a growing interest in investigating the existence of it in smaller organisms, some which do not exhibit spiking or a cortical structure, such as *c. elegans* (Leifer 2014). However, **there is a lack of evidence that predictive coding is used by *c. elegans*** to predict the onset of a sensory stimulus. It is important to know if smaller organisms exhibit predictive coding.

Convergent cross mapping (CCM) is a type of causal modeling method used in time series analysis to detect causal relationships between variables (Sugihara 1990). It is particularly useful for identifying nonlinear interactions **and can provide insights into whether one variable can predict another**. CCM **modelling is a way to identify if <u>neurons predict stimuli</u>,** it is used to generate future activity of a variable using past activity of another variable. The ability to make an accurate prediction about the future instance of a variable based on the time series generated manifold of another variable is what determines the degree of causal relationship between the two. I can use neural activity data to predict future sensory stimulus data, thereby proving that a neuron's activity predicts and anticipates sensory stimulus, the neuron is performing predictive coding.

<u>Hypothesis</u> **Predictive coding exists in simple organisms without hierarchical cortical structures, such as *c. elegans* and can be discovered using causal mapping techniques.**

<u>Significance</u>
Although predictive coding as a phenomenon has been characterized in more complex mammalian systems such as the human visual cortex, its presence has not been noted in less complex, non-mammalian, and non-spiking organisms. This can illuminate the predictive capabilities of less evolved organisms and uncover **novel simple circuits that perform complex computations such as predictive coding and will therefore offer valuable insights into this field.** It will offer insight into the basic building blocks of learning.

<u>Rationale</u>
I have calcium imaging data recorded from whole-brain activity of *Caenorhabditis elegans* (*C. elegans)* experiencing bacterial food stimuli. I have classified active neurons into six functional clusters: two sensory neuron clusters (ON, OFF), and four motor/command neuron clusters (AVA, RME, SMDD, SMDV). In my proposal I identify a computational method not commonly used to investigate this specific application of predictive coding: CCM. I further attempt to expand on this method by optimizing it, to recognize similar structural motifs regardless of magnitude within time series data. I will use this novel method to create a shadow manifold of the first half of calcium trace data from a neuron and use this to predict future stimulus, and compare this with the known future stimulus data. If the prediction is matches the real stimulus data closely, the causal relationship between the neurons and the stimulus is strong, and shows that the neuron can be used to predict the stimulus, thereby the neuron performs predictive coding.

Experiments
Aim 1 - Optimization of the Simplex Algorithm for Convergent Cross mapping (CCM)
Convergent Cross Mapping Model is a type of tool for short range forecasting of chaotic time series data. The algorithm it uses is called Simplex. Simplex works by reconstructing the system dynamics of a model based on Takens' Theorem. In which a single time series: $\{x_t, x_{t-1T}, x_{t-2T}, \ldots, x_{t-(E-1)T}\}$, where E is the dimensionality of a system, and T is some time lag tau. The original time series is the sequence for which we choose to make a prediction. This creates a shadow manifold. The forecast for the prediction can be denoted as: $Y(t_k) = \{Y(t_k), Y(t_{k-1}), \ldots, Y(t_{k-E+1})\}$ is made using the projections of its neighbors in the state space from the library set, $\{X_{(1)}, X_{(2)}, \ldots, X_{(E+1)}\}$, where $\|X_{(1)} - Y(t_k)\| = \min(\|X - Y(t_k)\|) = \min(\|X - Y(t_k)\|)$ for all X≠Y. So a time prediction one step forward into the future is $Y^{\wedge}(t_{k+1})$ is the average of all of the projections of all of its neighbors in this shadow manifold: $\{X_{(1)}(t_{1+1}), X_{(2)}(t_{2+1}), \ldots, X_{(E+1)}(t_{E+1+1})\}$ (Chang 2017). This method can be used to reconstruct manifolds which can predict the future activity of multiple variables in a single system, based on the time series of a single variable.

However, there are certain repetitive motifs within the time series data that are not utilized by the algorithm and may therefore lead to less than accurate inferences about the nature and strength of causal relationships among datasets. This may lead to a weaker predictive ability and therefore a weaker causal link between the two variables being studied. **I will create a modified version of this algorithm to consider repetitive motifs** of the same shape but various magnitudes. (Fig 1a) Shows calcium recording of two neurons in the *c. elegans* brain and the corresponding stimulus. One can see that although there are repetitive peaks with similar shapes, they may not be of the same magnitude. With the simplex algorithm as is, these points will lie in further positions on the shadow manifold, than they should be expected to. However, by normalizing each vector $\{X_{(1)}, X_{(2)}, \ldots, X_{(E+1)}\}$ according to its first term, this would account for these structural similarities. Given this modification does not aid in better predictive power, I will use the classic CCM.

Aim 2 - Establishing causal relationships among neuronal subtypes within *c. elegans* time series dataset using modified Simplex Algorithm
In Figure 1, there is an outline of the analysis procedure that one normally does using the current CCM method. I anticipate having better predictive capabilities with my altered simplex algorithm; however the workflow will remain unchanged. I will use both the original and the modified version and compare their predictive capabilities. Once the data (Fig 1a) is cleaned and normalized, one time series from a neuronal calcium response to a stimulus will be chosen. Simplex creates a shadow manifold which predicts the future time points of a time series, one half of the data is withheld as ground truth to test these predictions (Fig 1b). Based on trial and error a dimensionality E will be chosen. This is the dimensionality that yields the highest Pearson correlation between actual data and predictions. Multiple variables can then be tested, using CCM the causal relationship between one variable is determined based on the accuracy of the predictions that it makes for another variable's future activity. (Fig 1c) shows how the SMDV motor neuron predicts the AVA motor neuron, blue line, with the y-axis indicating Pearson coefficient between prediction and ground truth. The orange line shows the inverse condition, indicating SMDV has a stronger causal influence of AVA than the other way around. The algorithm can be run on an entire dataset of neurons (Fig 1d), and causal relationships can be between each variable within the whole system. **In my experiment I will do this between each neuron and the stimulus.** Fig 1e demonstrates the degree to which using this method to infer future activity of a neuron is better than using autocorrelation within the data. The time point at which the method is the most useful is beyond 150 time points into the future.

Possible Outcomes- I anticipate that when I run this script between a neuron and the stimulus I will **find a causal link between a single neuron and the stimulus time series itself. I will predict the future stimulus using past neural response. The optimal prediction range will be different for each neural subtype.** This would show that the **neuron is predicting the stimulus and is performing predictive coding**.

Alternative Approaches- Testing additional models to investigate the optimal model for the system
There is an abundance of other models which seek to predict behavior, and thereby establish a causal relationship between two variables throughout the course of time. Models, such Granger causality is a classic method in the field that is a good benchmark to test my modified CCM method on. Others, such as mixtures of structural causal models (MSCD) proposed by (Varambally 2023) rely on the use of multiple predictive parameters in a single system to explain its causal dynamics. It would be wise to different causal methods to explore the predictive capabilities of *c elegans* neurons, in case CCM does not offer strong predictive capabilities.
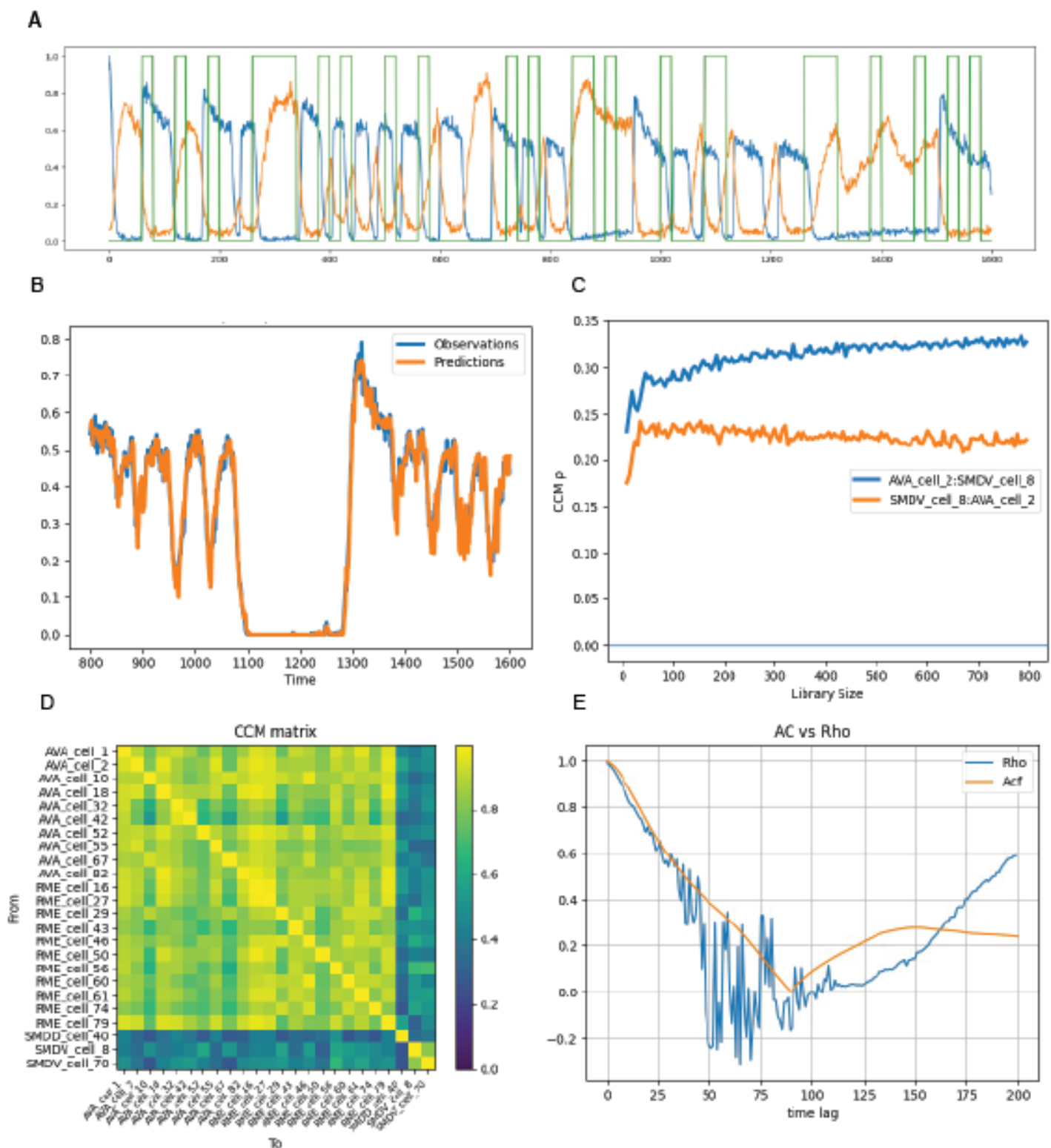
Figure 1 A) Time series calcium trace data from whole brain of c. elegans. Two traces (blue and orange) and the stimulus (green). B) Predictions (organge) created using the simplex algorithm of a single neuron's activity pattern, one 800 time points into with the future, using preciding 800 time points, with an embedding dimension size of 5, and a Tau prediction interval of 1. Overlayed against ground truthdata observations (blue). C) Convergent Crossmaping results showing the strength of the causal relationship based on predictive power between 2 neurons, AVA and SMDV. SMDV has better predictive power onto AVA than vice versa. D) Convergent Cross Mapping matrix of a series of AVA, RME, SMDD, and SMDV motor neurons in a single c. elegana shows strongest causal relationships between AVA and RME. E) Predicitive capability of the autocorrelation (orange) versus predicitive capability using simplex (blue) method shows advantage of using simplex past 160 time points.

References

1. Chang, C.-W., Ushio, M., & Hsieh, C.-H. (2017). Empirical dynamic modeling for beginners. *Ecological Research, 32*(6), 785-796. https://doi.org/10.1007/s11284-017-1469-9

2. Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences, 360*(1456), 815–836. https://doi.org/10.1098/rstb.2005.1622

3. Golkar, S., Tesileanu, T., Bahroun, Y., Sengupta, A. M., & Chklovskii, D. B. (2022). Constrained predictive coding as a biologically plausible model of the cortical hierarchy. In S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, & A. Oh (Eds.), *Advances in Neural Information Processing Systems* (pp. 14155-14169). Curran Associates, Inc. https://doi.org/10.48550/arXiv.2210.15752

4. Golkar, S., Tesileanu, T., Bahroun, Y., Sengupta, A. M., & Chklovskii, D. B. (2023). CUTS+: High-dimensional causal discovery from irregular time-series. *arXiv*.

5. Leifer, A. M., Fang-Yen, C., Gershow, M., Alkema, M. J., & Samuel, A. D. (2014). Optogenetic manipulation of neural activity in freely moving Caenorhabditis elegans. *Nature Methods*, 8(2), 147-152. https://doi.org/10.1038/nmeth.1554

6. Millidge, B., Seth, A., & Buckley, C. L. (2021). Predictive coding: A theoretical and experimental review. *arXiv*. https://arxiv.org/abs/2107.12979

7. Palmer, S. E., Marre, O., Berry, M. J., II, & Bialek, W. (2015). Predictive information in a sensory population. *Proceedings of the National Academy of Sciences, 112*(22), 6908–6913. https://doi.org/10.1073/pnas.1506855112

8. Sugihara, G., & May, R. M. (1990). Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. *Nature, 344*(6268), 734-741. https://doi.org/10.1038/344734a0

9. Varambally, S., Ma, Y.-A., & Yu, R. (2023). Discovering mixtures of structural causal models from time series data. *arXiv*. https://doi.org/10.48550/arXiv.2310.06312