The Use of Cross-Correlation Mapping in Identifying Backwards Projecting Connections between Visual Cortical Areas

Zack Cecere
UCSD

Abstract

The number of backwards projecting connections in the visual stream dwarfs the number of forward projecting connections. Yet, we have little understanding of the function of these connections. The primary obstruction to our understanding has been our inability to distinguish between the effect of feed-forward connections and feedback connections on a given neuron. In this work, I propose Cross-Correlation Mapping (CCM) as a means of isolating the effect feedback connections. Cross-Correlation Mapping is a time-embedding manifold method that has been successfully used to show unidirectional causality between two variables in complex dynamical systems [5]. I will, first, run CCM on model neural datafor which we know the causal relationshipsto determine the effectiveness of CCM. If I can show that CCM can uncover functional connectivity in neural populations, I will use CCM and fMRI data to create a functional mapping of the feedback connections from V2 onto V1.

1 Introduction

The main barrier to understanding feedback connections–like the ones traveling from V2 to V1–is the difficulty of isolating the effects of feedback connections from those of other network connections. This problem is equivalent to figuring out the effect of one input connection on a neuron while many other connections are also affecting the neuron's activity. This obviously difficult problem becomes even harder when we consider that there is no way to probe the activity of the majority of neurons feeding input into the target neuron. Cross-Correlation Mapping–a method developed in George Sugihara's lab at Scripps Oceanographic Institute–has been successfully used to determine the causal relationship between two variables in a high-dimensional, dynamical system [5]. If we think of individual neurons as variables in a high-dimensional system–a neural network–we may guess that CCM could be used to decipher brain architecture.

2 Methods

2.1 Cross-Correlation Mapping

According to dynamical systems theory, two time-series variables are causally linked if they are from the same dynamic system [5]. For example, variables X and Y in the Lorenz attractor are causally linked because the state of X is dependent upon the previous state of Y and vice versa. We can also think of dynamical causality between two variables in terms of their shared attractor manifold. Since X is contained in the state determining equation for $\frac{dY}{dt}$ and Y in the state determining equation for $\frac{dX}{dt}$ in the Lorenz attractor, we know, given no change in any other variable in the system, that the X, Y phase plane will contain a coherent, well-defined attractor. If X and Y do not exhibit a causal

relationship, no attractor would exist–nor a phase plane, really–as the quantities $\frac{dX}{dY}$ and $\frac{dY}{dX}$ are undefined.

The technique pioneered by Sugihara et al [5]–Cross-Correlation Mapping (CCM)–uses the dynamical systems definition of causality. More specifically, CCM constructs shadow manifold for two variables in the proposed system using time-lagged coordinate embedding. That is, CCM creates a shadow manifold for each variable independently. CCM creates a given point on the shadow manifold by sampling an individual variable multiple times with a fixed interval. It then checks whether nearby points in one shadow manifold correspond with nearby points in the other. There are many points that have a similar value in the parent manifold–the original manifold without any embedding–for a single variable but are not actually close to one another–that is, they have different values for the other variables. The time-lagged embedding seeks to distance these points: points not near each other on the parent manifold are unlikely to evolve through time similarly. If contiguous points in one shadow manifold can be used to identify contiguous points in another, we can build a section of a coherent higher-dimensional manifold from these contiguous points–that is, we can theoretically rebuild parts of the parent manifold from the relationship between the shadow manifolds. Thus, CCM works because a well-defined relationship between shadow manifolds indicates that a coherent attractor manifold exists in a space containing the two variables.

In CCM, the cross mapping between the shadow manifolds mentioned above is done for a number of library sizes. The library size is the number of points–not the dimension–used to build the shadow manifold. If the two variables related to the shadow manifolds are causally related, cross-mapping should improve with increased library size. This is because, with increased library size, there is a higher probability that the points identified as contiguous in one of the shadow manifolds actually are contiguous. Thus, the degree of convergence is the best measurement of causality in CCM analyses.

2.2 Jumbled CCM

The canonical methods for fMRI processing [6] output a scalar response value for each voxel to each image. Thus, we are no longer dealing with time series in fMRI data. There remains no ordering to the series as image 1 could induce a more similar overall response to image 50 than image 2. I created the jumbled CCM method to test whether the loss of a natural ordering in the data series affects CCM. Essentially, jumbled CCM works by repeatedly applying CCM while shuffling the input data. The idea is, after enough shuffles, certain points in the shadow manifold will exhibit consistent contiguity. Thus, through averaging, CCM should generate an internal representation of the true shadow manifold, despite us not actually knowing the true form of this manifold.

2.3 False Neighbors Analysis

The minimal number of dimensions–number of time lagged values of a variable that goes into each point on the shadow manifold–that allow for the separation of points that are not near each other on the original manifold–false neighbors–is the ideal dimensionality for a time series embedding. In their 1992 paper [3], Kennel et al devised a method for finding this ideal embedding dimensionality. They hypothesized that false neighbors would distance themselves given an increase in embedding dimension.

$$\left(\frac{R_{d+1}^2(n,r) - R_d^2(n,r)}{R_d^2(n,r)}\right) > R_{tol} \tag{1}$$

Where R_d is the euclidean distance between a given point and its nearest neighbor for dimensionality d.

Kennel et al [3] noticed that this metric failed to parse out false neighbors with noisy data. They realized that noise causes the radius of the original manifold to increase to such a degree that the distance between every pair of points in the embedded manifold for a given series is very large. In order to account for this, they devised a metric that compares the distance between two points in a given embedding dimension to the radius of the attractor.

$$\frac{R_{d+1}(n)}{R_A} > A_{tol} \tag{2}$$

The idea is that false neighbors will be kicked out to the edge of the attractor by adding additional embedding dimensions.

Typically, the method of false neighbors outputs, as one would expect, the number of nearest neighbors found to be false by increasing the dimensionality by one. In this study, I simply calculate the values of metrics 1 and 2 averaged across a number of points for a number of dimensions, assuming that these metrics are proportional to the number of false neighbors.

2.4 Granger Causality

Granger causality, in its simplest form, uses linear-autoregression to determine whether the value of one variable can be predicted from another.

$$X_1(t) = \sum_{j=1}^{p} A_{11,j} X_1(t-j) + \sum_{j=1}^{p} A_{12,j} X_2(t-j) + E_1(t) \tag{3}$$

In the above example, variable X2 Granger-causes variable X1 if the variance of E1 is reduced by the inclusion of X2 in the first term [2]. When the variance of the error term E1 decreases in this way, we know that we can use X2 to better predict the statistical structure of X1. Granger causality and CCM are similar in the respect that they measure how well the statistical structure of one variable can be used to predict that of another. The primary difference between the two methods is that CCM is tuned for and gives more information about any manifold that may pass through the space encompassing the two variables.

We implemented two forms of Granger causality. Our first-order implementation simply uses the responses of a variable number of voxels to about 1000 natural images to predict the behavior of another voxel to the same images using least-squares optimization. This method then stores the variance of the prediction error. We also implemented a second-order Granger causality method that considers correlations between the input voxels in addition to the raw values of the input voxels. The second-order method exhibited superior performance but proved to be too slow to use–unparallelized–for large data sets.

3 Results

3.1 Testing CCM on Chaotic Data

In order to determine whether CCM works at all, we tested the method on several chaotic systems. First, we used a simple, two-state chaotic system with variable causality strength between the two variables.

Since the CCM result converges for the non-zero causality case and not for the zero causality case, we know that CCM is able to identify causality for simpler chaotic systems.

Sugihara et al [5] claimed that CCM could determine the direction of causality between two variables. We tested this claim using a modified form of the previously mentioned chaotic model. In this new version of the chaotic model, variable X has a causal effect on variable Y but variable Y does not have a causal effect on variable X.

The contrasting results in figures 2 and 3 indicate that CCM is, in fact, able to determine the direction of causality between two variables.

Most interesting systems have many interacting variables. In order for CCM to be useful, it must have the ability to identify causality between pairs of variables in these complex systems. We tested CCM's ability to find causal relationships in these complex systems using a 5-state chaotic model in which only certain variables are causally related.

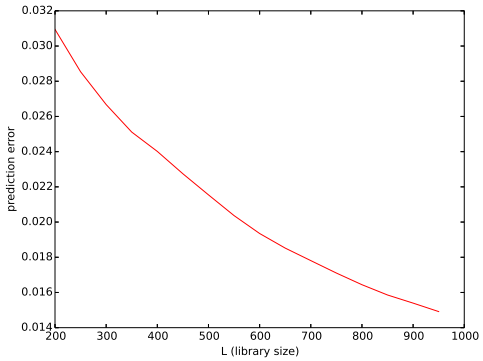


Figure 1: Results of a CCM run on two variables in a chaotic model known to exhibit dynamic causality. The improvement of prediction with library size indicates that CCM correctly identifies this causal relationship.

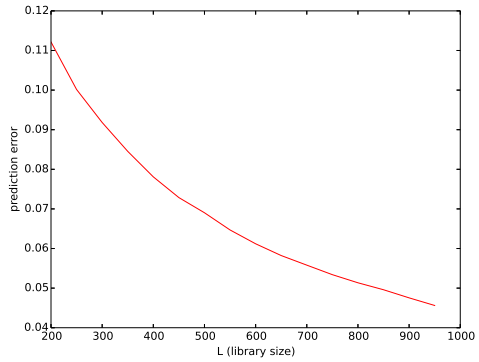


Figure 2: CCM test in the direction of X$--$ $>$Y. The improving performance of CCM with library size shows that CCM correctly identifies X as causing the behavior of Y.

CCM correctly identifies variables 2 and 3 as causing the behavior of variable 1. It also correctly shows–via a lack of improvement in correlation index (inverse of prediction error)–that variable 4 does not cause the behavior of variable 1. This result is very encouraging as it allows us to apply CCM to high dimensional neural systems.

3.2 Testing CCM on Hodgkin-Huxley Networks

Given that we can describe most neural systems with dynamical equations, we expect to be able to apply CCM to neural data. More specifically, we wish to apply CCM on the level of neurons and neuronal populations, finding causally related–aka, functionally connected–neurons and neuronal populations. To test whether we can apply CCM in this way, we generated the following sample feed-forward model composed of Hodgkin-Huxley neurons:

Several interesting observation arose in the analysis of this network. First, we notice that CCM identifies neuron 0 as being functionally connected to neurons 2, 5, and 6–evidenced by the convergence in figures 5, 6, and 7.

Although neuron 0 is technically only connected to neuron 2, the fact that CCM identifies it as being functionally connected to neurons 5 and 6 is both useful and correct. Neuron 0, through neuron 2, singularly controls the behavior of neuron 5, and controls the behavior of neuron 6 in concert with neuron 1. Moreover, the CCM output converges to the lowest value in the case of the 0 $--$ $>$ 2

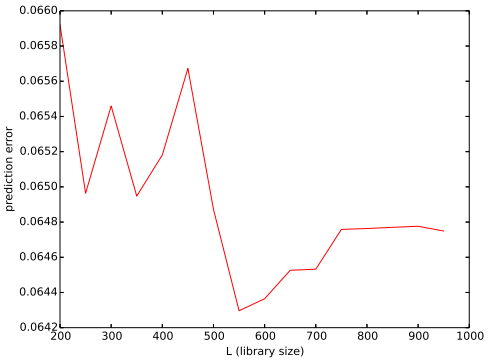
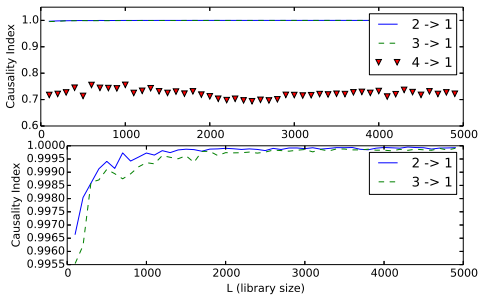


Figure 3: CCM test in the direction of Y$--$ $>$X. One could make the claim that this figure shows convergence. However, the level of improvement with library size, in this case, is not above chance, indicating the CCM correctly shows there to be no causal effect of Y on X.



Testing the effectiveness of CCM on chaotic model where variable 1 causally affects variables 2 and 3, but not variable 4. CCM correctly recognizes these causalities as only the 2->1 and 3->1 connections' causality indices converge to 1. Note: the second plot is a zoomed in version of the first.

connection, providing us with a means of separating the tiers of connectivity. Thus, in the case of neural data, the consistently decreasing / converging behavior of CCM output identifies functional connectivity–further evidenced by the lack of convergence in the 0 $--$ $>$ 4 CCM run–while the values at convergence allow for the ranking of connectivities in a network.

In our testing of various network architectures, we found that CCM struggles to identify only one connection type. If a given Hodgkin-Huxley neuron receives an input from only one other neuron, CCM cannot not deduce whether the connection projects onto the given neuron or from the given neuron. For example, CCM was unable to determine the directionality of the connection between neurons 0 and 2 in the above network.

Although CCM suggests that neuron 0 synapses onto neuron 2 and vice versa, it performs much better–in terms of noise of the curve and converged prediction error–in the correct direction–the 0 to 2 direction. This sparsely connected case will likely not matter in real neural data because real neural networks are much more highly connected and noisy than our Hodgkin-Huxley simulations. However, this failure of CCM requires us to apply CCM in both directions for a hypothetical connection in order for us to be certain of the directionality of the connection.

Thus far, we have demonstrated that we can easily apply CCM on the level of small networks. However, CCM would prove even more useful if we could apply it on the level of large-scale brain dynamics. To this end, we will attempt to use CCM in conjunction with fMRI data to gain some insight on large-scale connectivity in the visual cortex.

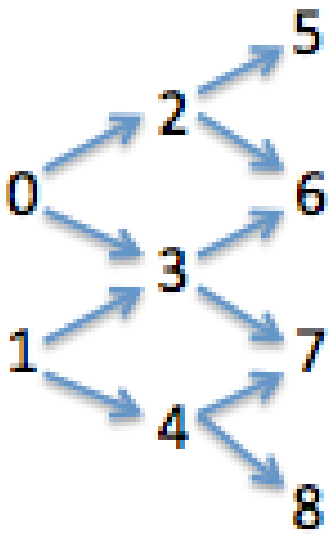


Figure 4: The network used to generate the Hodgkin-Huxley data.

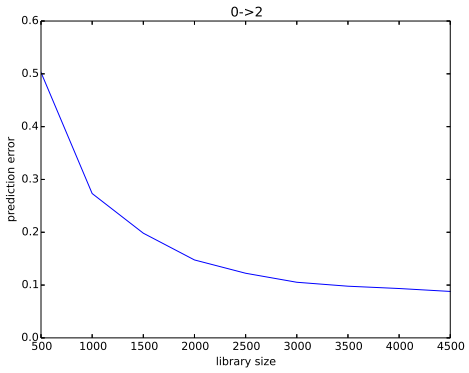


Figure 5: CCM correctly identifies a functional connection from neuron 0 to neuron 2.

3.3 Applying CCM to fMRI Data

Thus far, we have demonstrated that we can apply CCM on the level of small networks. However, CCM would prove even more useful if we could apply it on the level of large-scale brain dynamics. To this end, we will attempt to use CCM in conjunction with fMRI data to gain some insight on large-scale connectivity in the visual cortex.

We obtained visual cortical fMRI data from the Gallant lab [6]. The Gallant lab generated the data by taking fMRI reading from humans viewing natural scenes. The response of each voxel to each image was determined using the hemodynamic response convolution protocol discussed in [6]. Since the protocol assumes that each voxel has a scalar response–more specifically, a scale on the unchanging HRF of the voxel–to a given image, it outputs data that is not a time series. To deal with this unordered data, we created the jumbled CCM method.

Before applying the jumbled CCM method to fMRI data, we tested the method's validity on the chaotic and Hodgkin-Huxley data already discussed.

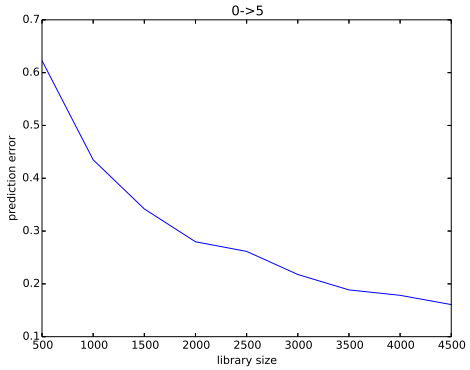


Figure 6: CCM correctly identifies a functional connection from neuron 0 to neuron 5, despite the fact that neuron 5 is two edge lengths away.

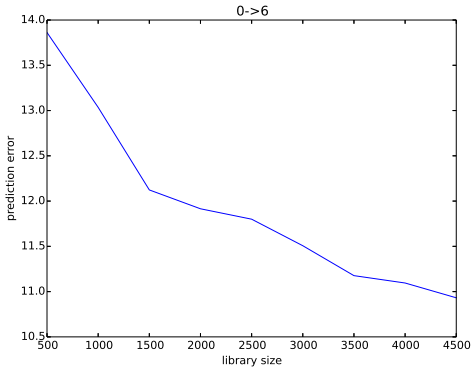


Figure 7: CCM correctly identifies a functional connection from neuron 0 to neuron 6. This is an especially encouraging result as neuron 6 receives information from both neuron 0 and neuron 1–two neurons with very different activities.

In the methods section, we hypothesized that the jumbled CCM method could generate an internal representation of the shadow manifolds of the two input variables, despite us not knowing the true from of these manifolds. The success of the jumbled CCM in regenerating our earlier data–both the chaotic and Hodgkin-Huxley models (data not shown)–indicates that this hypothesis has merit.

Before moving on to the fMRI analysis, we needed to determine the embedding dimension our CCM method should use. As discussed in the methods section, using too low of an embedding dimension will often produce false neighbors. We used the method of false neighbors to determine the optimal embedding dimension [3].

After applying the method of false neighbors to voxel 100 of V1 in the fMRI data–the central voxel of our study–we decided to use an embedding dimension of 70.

Due to the many averaging steps required in jumbled CCM, the method is very slow. Thus, we decided that, for the sake of time (and the fact that I ran all of the simulations on a MacBook Pro) that we would analyze how the behavior of only voxel 100 in V1 (primary visual cortex) is controlled by voxels in V2.

In figure 13, we see that a small region of V2–from about voxel 200 to voxel 300–exhibits a compressed range of values to which the CCM prediction error converges. This trend becomes more obvious when we average the values of nearby voxels (figure 14). Thus, there exists some indication that this area of V2 project backs to voxel 100 in V1.

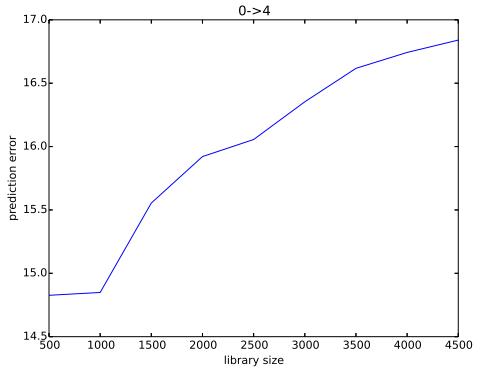


Figure 8: CCM correctly identifies the lack of a functional connection from neuron 0 to neuron 4.

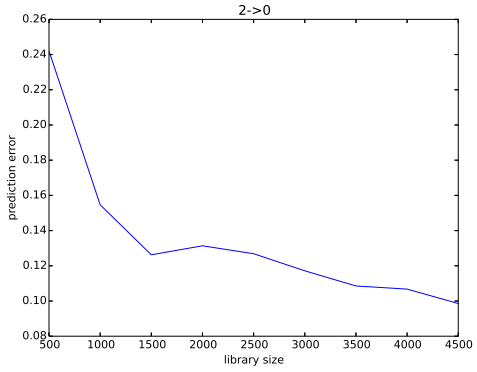


Figure 9: CCM falsely finds a connection from neuron 2 to neuron 0. This is a result of neuron 2's behavior being totally determined by the behavior of neuron 0.

In order to validate our hypothesis that voxels in the 200-300 region project back to voxel 100 in V1, we tested CCM exhaustively on voxels in this region. More specifically, we averaged thousands of jumbled CCM trials, attempting to eliminate as much noise as possible. We believed that this method was effective because every massively averaged CCM run on the same two voxels produced the same library size vs. prediction error figure. Figure 15 shows the massively averaged CCM output for the hypothetical connection from voxel 215 in V2 onto voxel 100 in V1.

Figure 15 strongly suggests that voxel 215 in V2 projects onto voxel 100 in V1. In fact, if we assume every change in prediction error due to the addition of new points for consideration has a 50:50 chance of being negative–as is the case in CCM analyses of random data (data not shown)–the binomial theorem indicates that there is less than a 5 percent chance of obtaining as many decreases with increasing library size as seen in figure 15. Coupling this analysis with the fact that many of the voxels near 215 show similar relationships with voxel 100 in V1, we believe that this region of V2 projects back to V1.

We discovered in our Hodgkin-Huxley analysis that CCM sometimes struggles in determining the directionality of a connection. Keeping this in mind, we ran CCM to determine whether voxel 100 in V1 projects onto voxel 215 in V2.

Figure 16 gives no indication of voxel 100 in V1 projecting onto voxel 215 in V2. This further supports our result: voxel 215, as well as other nearby voxels in V2, projects functional connections back to voxel 100 in V1.

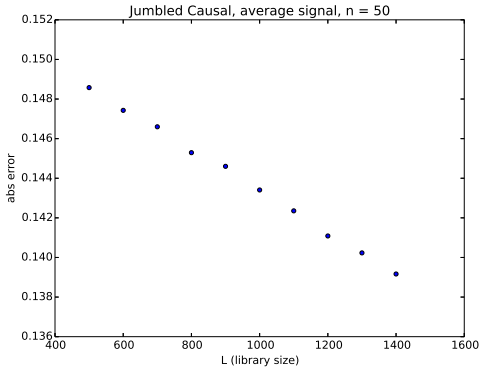


Figure 10: The jumbling does not interfere with CCM's ability to identify a causal connection; it still recognizes dynamical causality in this chaotic model.

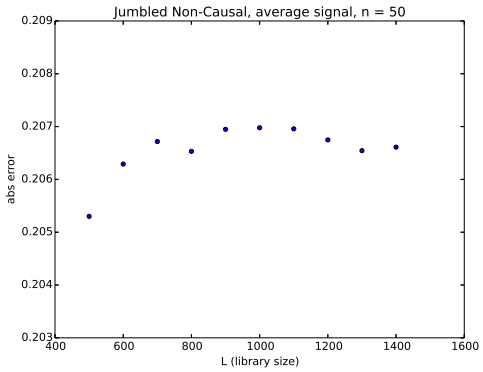


Figure 11: The jumbling does not interfere with CCM's ability to identify the lack of a causal connection. This was done for two un-related variables in a high-order, chaotic model.

Thus, our data recommends that the response of voxel 100 in V1–and likely many other voxels in V1–is influenced in a dynamically causal manner by a population of voxels in V2.

3.4 Comparison to Granger Causality

In the hope of amassing more evidence for functional backwards-projecting connections from V2 to V1, we implemented a form of Granger causality. More specifically, we used a first-order Granger causality method (Methods) to determine which V2 voxels best predict the behavior of voxel 100 in V1.

Our scan of V2 voxels that Granger cause' the behavior of voxel 100 in V1 produced similar–but certainly not identical–results to those of CCM. One interesting result of the Granger causality scan is that, on average, the voxels around voxel 215 of V2 predicted the statistical properties of voxel 100 in V1 most accurately. This directly agrees with the CCM analysis. However, the Granger causality scan indicates that the V2 region around voxel 75 can also be used to accurately predict the behavior of voxel 100 in V1. The CCM analysis unearthed no such pattern.

4 Conclusion

Our Hodgkin-Huxley analysis shows that Cross-Correlation Mapping is a valid technique for determining functional connectivity in small neural networks. Moreover, the fact that CCM performance

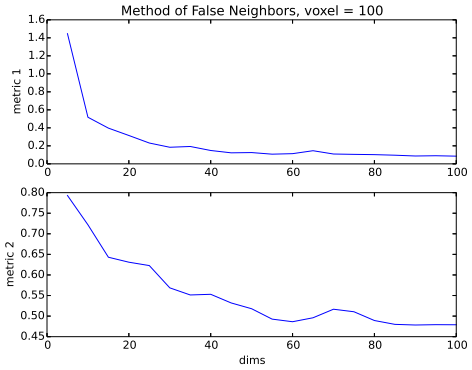


Figure 12: Results of method of false neighbor analysis on voxel 100 in V1 of the fMRI data. The y-axis in the top figure is equation 1, while the y-axis on the bottom figure is equation 2 (methods). Together, the figures indicate that we cease to gain a benefit–root out false neighbors–from increased embedding dimensionality at about 65 dimensions.

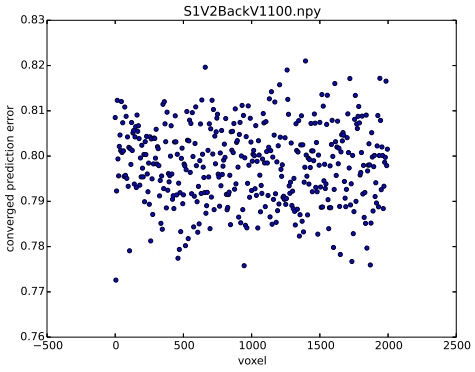


Figure 13: Results of running CCM to determine which voxels in V2 (V2 voxel number shown on x-axis) causally effect the behavior of voxel 100 in V1. The y-axis is the value to which the CCM run converged. We used converged value for this figure because the value to which a CCM run converges is stable between CCM runs. The initial behavior of the CCM output, however, requires the averaging of many CCM runs to gain a stable shape. Thus, we could only analyze the initial behavior for a small number of voxels.

actually improves with more connectivity and noise suggest that CCM can be applied to all types of neural data. Indeed, our application of CCM to visual cortical fMRI data provided very strong evidence for the existence of backpropagating information. This finding is a very nice support for the predictive coding model of visual processing [4].

Although we achieved some success in our use of CCM, the jumbling method we were forced to use in our fMRI analysis was both slow and noise inducing. We are currently working on applying the maximum noise entropy method [1] to find the visual features to which each voxel in a number of visual areas responds. We will then apply these voxel response models to a number of slowly morphing visual stimuli to obtain intrinsically ordered data–that is, stimulus 1 will be more similar to stimulus 2 than to stimulus 50. This method should grant us a representation of the shadow manifolds without the need for jumbling.

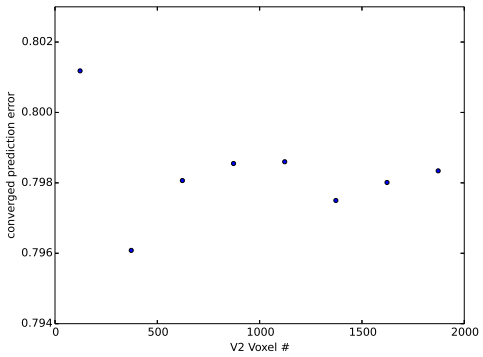


Figure 14: What happens after we average the converged values of nearby voxels (essentially, binned the results of figure 12). Here, we see more clearly the dip in CCM converged value in the 200-300 voxel section.

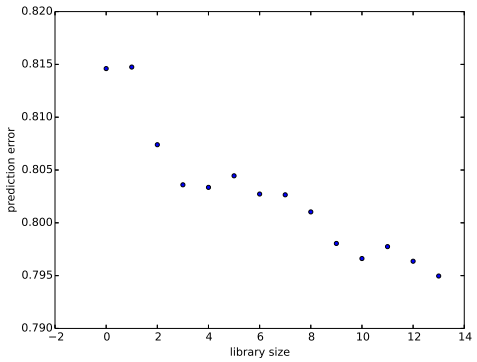


Figure 15: The signal produced by the averaging of many CCM runs testing whether voxel 215 in V2 functionally causes the behavior of voxel 100 in V1. The fact that pretty much every set of averaged CCM runs for this voxel combination converges to this improving estimation pattern strongly advocates for a function connection from voxel 215 in V2 to voxel 100 in V1.

5 References

[1] Fitzgerald et al. Second Order Dimensionality Reducton Using Minmum and Maximum Mutual Information Models. PLoS Computational Biology 7. (2011).

[2] Granger, C. W. J. Investigating causal relations by econometric models and cross-spectral methods. Econometrica 37. 424-438 (1969).

[3] Kennel, Brown, and Abarbanel. Determining embedding dimension for phase-space reconstruction using geometrical construction. Physical Review A 45. (1992).

[4] Rao and Ballard. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nature Neuroscience 2. (1999).

[5] Sugihara et al. Detecting Causality in Complex Ecosystems. Science 338. (2012).

[6] Kay et al. Identifying natural images from human brain activity. Nature 452. (2008).

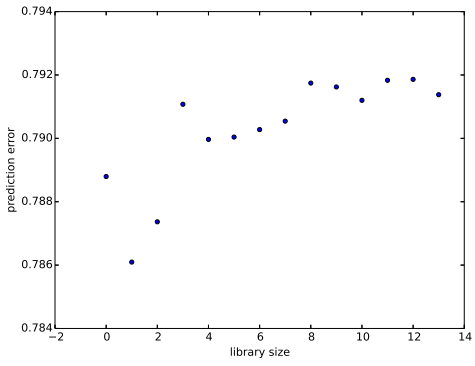


Figure 16: We used the same averaging of CCM runs technique here that we used for figure 14. The loss of performance with increasing library size indicates that there is no little causal connectivity from voxel 100 in V1 to voxel 215 in V2.

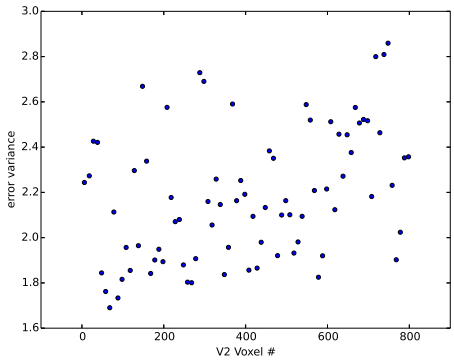


Figure 17: Each point corresponds with a prediction of the behavior of voxel 100 in V2 from 30 voxels in V2 starting at the x-value of the point.

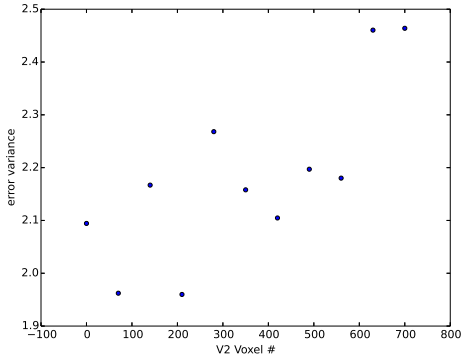


Figure 18: This figure was generated by binning the previous figure.