1 Introduction

Human beings have an insatiable desire to understand, not only the things around us, but also ourselves. Why do we behave and respond the way we do [12]? How do we learn and adapt [11][10]? What do we believe or value [22]? These questions have mesmerized wise men for thousands of years, and the answers continue to evade us. For many of us, the key to all the secrets about ourselves is our brain, the central hub of all our thoughts and decision making. So, when a new technology emerged allowing us to study the brain in a quantitative manner, it opened countless doors of opportunities for those who possessed an aptitude for quantitative analysis and a hunger to learn more about our own identity.

Functional Magnetic Resonance Imaging, often referred to as fMRI, was first discovered and applied as a brain mapping method in 1990 by Seiji Ogawa [25]. Since then, this technology quickly popularized among Brain Science related research due to its unprecedented low health risk to subjects and its ability to accurately translate brain activities into highly-structured data [21]. Once the the highly abstract brain activities are converted into signs and digits, decoding patterns in the brain becomes a possibility. By treating each fMRI brain image as a feature vector, machine learning algorithms trained on a subset of the images may be used to distinguish cognitive information (e.g. which part of a movie someone is watching) from the held-out images [26][17][18]. These approaches typically treat each brain image in isolation, and attempt to identify patterns of activity associated with each of several candidate brain states. However, measurements from fMRI are often corrupted by noises created by the environment, human error, random neural activity, etc [16]. To address this issue, new branches of research has been emerging that specifically focuses on the extraction of useful information from noisy brain fMRI data by treating the brain images as a dynamic time series, and one of the most fruitful branches is functional connectivity [14] [19] [20].

Functional connectivity analyses entail computing the correlations between the time series of activations each pair of brain regions exhibits. When we observe the brain, the neural activities can appear incredibly complex. But mechanics of brain are far from random: rather, the dynamic patterns of activity our brains exhibit are highly structured. Pre-

sumably this mirrors the complex but highly structured nature of our internal thoughts and experiences. Recently, it has become clear that important cognitive information is contained in higher-order brain patterns, such as the dynamic correlational structure of the data [6]. When applying functional connectivity analyses on one subject in isolation, one can examine how (or whether) the correlational structure of these activity patterns (across a given set of brain regions) varies according to the cognitive task an experimental subject is performing [27][28]peterson10. Alternatively, the inter-subject functional correlation (ISFC) analyses applies functionally connectivity analyses across multiple subjects to isolate stimulus-dependent inter-regional correlations from intrinsic neural processes and non-neuronal noise [5][1].

Over the past decade, functional connectivity analyses of brain data has evolved significantly [4][8]. However, there are three fundamental limitations to past approaches:

1. **Correlation matrices are not scalable.** Examining pairwise correlation in brain data produces a correlation matrix with O(n2) entries (where n is the number of brain regions). When n is large (e.g. the number of brain regions in an fMRI volume), the full correlation matrix can become unwieldy to compute with [28][29][30][31]. Furthermore, if one wishes to examine higher-order patterns (e.g. how correlations between correlations change over time), the storage requirements of the resulting patterns increase exponentially.

2. **The sliding window approach is not well suited to studying dynamic activity.** Most approaches to calculating functional connectivity uses the sliding window method, where a window of set time length is selected to calculate the functional connectivity at one time point before the window is shift forward for following calculations [3][2]. One disadvantage of the sliding window approach is the loss of information for a number of time points equal to the sliding window length. Although many practical applications use buffers to make up for the loss, repeated application of the sliding window approach on the same dataset——e.g. calculating the correlation of the correlation between nodes——is impractical. In addition, the sliding window approach provides only a poor approximation of the moment-by-moment patterns at the heart of these representations.

3. **Only using node activations and first level dynamic patterns may not be enough.** The brain is a network of seemingly autonomous nodes that are actually intricately interconnected. Useful information about the brain may exist within interactions between interactions between nodes (2nd order), interactions between interactions between interactions (3rd order), or in even higher order dynamics. To fully grasp the functions of a single node in the network, it is crucial to understand of its activities relative to the richly woven network that supports it. Therefore, incorporating information from higher order dynamics could potentially improve the quality of brain analysis.

Here, we present the High-Order Brain Dynamics (HOBD), a model that seeks to understand the dynamics behind different "levels" of interaction between brain structures and how this information is represented in the brain. To efficiently access and analyze brain dynamics at higher levels, the HOBD model revolutionizes functional connectivity calculation methods and creates a means to repeatedly "level up" brain dynamics information, computing higher level functional connectivity from low level information. These changes have notably improved the recovery accuracy of brain correlation structures, and significantly increased decoding capabilities using a limited amount brain data.

First, to find an effective replacement for the widely used sliding window method, we designed TimeCorr, an intuitive method that found inspiration directly from the fundamental correlation function. Like the sliding window method, TimeCorr is able is able to recover functional connectivity from fMRI brain images with similar if not higher accuracy. But TimeCorr goes beyond the sliding window method by (a) not requiring buffer at each end for calculation, thereby avoiding data loss, (b) offering extra stability by using all the time points in the time series for calculation of correlation at every time point, (c) provides the option to shift between locality and fluidity based on user demand by varying the resolution parameter.

TimeCorr achieves all the above functions through the use of Gaussian distribution. By attaching a Gaussian coefficient to each time point component in the time series, TimeCorr is able to allocate the amount of influence each neighboring time point has on the calculation of correlation at the time point of interest. As the Gaussian center (highest density/coefficient)

is always at the time point of interest, TimeCorr guarantees to return highly accurate approximations of the temporal ground truth. In addition, the user can choose to have more fluidity in overall results by choosing coefficients from a Gaussian distribution with a large variance; or higher resolution and more locality by choosing coefficients from a distribution with a smaller variance.

Second, we designed a "level up" method that applies TimeCorr, Inter-Subject Functional Connectivity (ISFC) and Principle Component Analysis (PCA) to calculate higher order functional connectivity. To overcome the problem of scaling and to enforce uniformity for inter-levels analysis, we use TimeCorr on each subject to calculate functional connectivity from previous level activations and then reduce the results to an arbitrary number of feature points to represent the activations for the next level. As TimeCorr avoids data loss in calculating functional connectivity and PCA maintains an uniform number of activations at each "level up", we are able to obtain 10^{th} order functional connectivity while ensuring linear scaling in storage and computation.

We believe that different orders of functional connectivity represents information on different aspects of brain dynamics, and incorporating knowledge of the higher-order structures of the brain will expand present brain decoding capabilities. To measure the amount of information present at each level, we introduce decoding accuracy, a testing parameter that describes the proportion of time points in one level of a subject's functional connectivity graph that has the highest correlation with the same time point from the average activations of all other subjects. We are interested in how decoding accuracy changes as we reach higher order functional connectivity. Although we presumed brain dynamics would converge as we get to higher orders and decoding accuracy would increase, reality proved otherwise. This is shown and discussed in more detail in the Intra-Level Decoding Analysis section of our results.

To carry out an in-depth evaluation of the effectiveness of the HOBD model, we also conducted the Multi-Level Mixture Analysis. This study aims to understand how the dynamic information of different levels of interactions between brain regions can be used to improve brain pattern decoding capabilities. In the Multi-Level Mixture Analysis section of our results, we provide a side by side comparison that highlights the improvements achieved by

a mixture model that incorporates multiple levels of functional connectivity. In the process of finding the mixture model that maximizes decoding accuracy, we also explored the utility of each level of brain dynamics through analysis of the distribution of optimal level weights. The methods and findings will be discussed extensively in the results section.

2 Model Description

The High-Order Brain Dynamics model seeks a new perspective to understanding the dynamics behind stimulus-driven brain activities. Utilizing the capabilities and advantages of TimeCorr and "Level Up", the HOBD model attempts to incorporate information on high-order brain dynamics into its analysis of the brain.

2.1 Single Subject TimeCorr

The TimeCorr method was designed for the purpose of replacing the sliding window method as a lossless alternative to achieve more accurate calculation of functional connectivity (the correlation structure between brain regions within a subject's brain) from brain fMRI data. The sliding window method operates by applying the correlation calculation function on a block of data centering on the time point of interest t and spanning L_w time points, where L_w is the sliding window length and must be of considerable magnitude for the sliding window method to achieve respectable accuracy. However, due to its inherent design flaw, using the sliding window method on a dataset of time length T could only yield functional connectivity for $T - L_w$ data points, a loss not insignificant due to the typical magnitude of L_w. Thus, as calculating high order functional connectivity involves repeatedly applying the method of correlation calculation on the dataset, using the sliding window method would result in rapid deterioration in the amount of usable data. In addition, as the sliding window method uses a long block of data for correlation calculation, it can only achieve a rough approximation of the general average in the block and not the instantaneous truth at each time point.

To solve these problems, we designed the TimeCorr method, which finds inspiration from the fundamental correlation function. Instead of isolating a block of points for correlation calculation, TimeCorr effectively utilizes information from the entire dataset for functional connectivity calculation at every time point. To ensure locality of TimeCorr's results, every term is multiplied by a weight drawn from a normalized Gaussian density function so that each time point influences the calculation of functional connectivity at time point of interest t proportional to its distance from t. This adaptation of the correlation function effectively

eliminates results repetition at the beginning and end of the time series, thereby increasing the number of output time points to equal to the number of inputs and avoiding any data loss. In addition, through the application of Gaussian density coefficients, TimeCorr gives more significance to information closer to the time point of interest in its calculations, thereby ensuring a more accurate estimation of instantaneous correlation.

Formal Definition

Given a single subject fMRI dataset with T time points and N brain regions, to apply TimeCorr:

1. Base on the amount of influence neighboring time points should have on the calculation of the functional connectivity at each time point, choose variance V for the Gaussian density function to generate coefficients.
2. For each time point t:
   (a) Using a Gaussian density function of variance V and mean t, generate an array of coefficients w_t of length T.
   (b) For each brain region, element-wise multiply its activation time sequence a_i by the coefficients array to create a weighted activations array S_t^i.
   (c) Find the temporal correlation between brain region i and brain region j at time point t through the equation:

$$C(S_t^i, S_t^j) = \frac{1}{Z} \frac{\sum_{l=0}^{T} (S_l^i - \bar{S_t^i}) \cdot (S_l^j - \bar{S_t^j}) \cdot \mathcal{N}(l|t, \sigma)}{\sigma_{S_t^i} \cdot \sigma_{S_t^j}}$$

Where

$$Z = \sum_{l=0}^{T} \mathcal{N}(l|t, \sigma)$$

$$\bar{S_t^i} = \frac{1}{Z} \sum_{l=0}^{T} S_l^i \cdot \mathcal{N}(l|t, \sigma)$$

$$\sigma_{S_t^i} = \sqrt{\frac{1}{Z} \sum_{l=0}^{T} (S_l^i - \bar{S_t^i})^2 \cdot \mathcal{N}(l|t, \sigma)}$$

(d) Repeat above process for every pair of brain regions to create correlation matrix for time point t

(e) Convert the correlation matrix at every time point to its inverse squareform format, output an array of size T by $(V^2 - V)/2$ dimensional matrix, where T and V represent the time length and brain region count of the dataset.

The addition of the Gaussian density variance parameter V gives the user the extra freedom to customize the desired level of resolution for correlation calculation. If a large variance is chosen, the influence of neighboring time points on the calculation of functional connectivity at each time point is increased, thereby adding smoothness and stability to the resulting functional connectivity time sequence and providing a more accurate representation of the general trend. But if a small variance is chose, most of the weight will be placed on the time point of interest and its immediate neighbors, and so the resulting functional connectivity time sequence will have more emphasis on local accuracy. After experimenting with different setups, we discovered that the best way to distribute weights for time points in the time series is to apply a Gaussian probability density function centered around the time point of interest with variance equal to the minimum between the total number of time points and 1000. This finding will be discussed in more detail in the Results section.

In contrast to the sliding window method, which is widely considered to be the golden standard for fMRI functional connectivity calculation, the TimeCorr approach is able to more accurately retrieve the temporal correlation at each time point without loss of important data. This advantage makes it possible to "level up" to higher order functional connectivities—progressively finding the correlation of correlations—and explore dynamics and patterns heretofore untouched. We will go into more detail about the "level up" process and high order functional connectivity in the "Leveling Up" section.

2.2 Inter-Subject Functional Correlations using TimeCorr

The Inter-Subject Functional Correlations (ISFC) is a process of calculating functional connectivity between brain regions of different subjects [1][5]. The patterns recovered by ISFC is analogous to single subject functional connectivity (which reflect the correlational

structure across brain regions within an individual's brain), but they should reflect only activity that is specifically stimulus-driven. For every subject, we compare its brain activations time sequence a_i with the average activations of all other subjects a_{-i}. Through the process of averaging activations of multiple subject, we dampen the effect of noise and enhance stimulus-dependent activations in a_{-i}. Therefore, when we calculate the functional connectivity matrix between subject activation a_i and the average response from the rest of the participants a_{-i}, the functional connectivity we calculate are more likely to reflect activities that are stimulus-driven.

Furthermore, after we obtain the functional connectivity matrix between each subject and the average of their counterparts, we use Fisher Z-transformation to average the results from all the subjects. In addition to the noise-dampening and stimulus-enhancing effects that comes with averaging, we are also able to gain an unbiased view of the overall connectivity pattern across all subjects within our analysis. Fisher Z-transformation is applied in the averaging process as a means to stabilize and reduce approximation variance to return a less biased result than from directly averaging correlations.

Formal Definition

Given a fMRI dataset of S subjects, each with T time points and N brain regions, to apply TimeCorr ISFC:

1. Determine the desired temporal resolution for calculation and select σ, the Gaussian density function variance for TimeCorr coefficients generation.
2. For each subject s,
   (a) Find the average activation of all other subjects:
   $$O_s = \frac{\sum_{i \neq s}^{N} S_i}{N-1}$$
   where S_i represents the activation matrix for subject i and N represents the total number of subjects.
   (b) Find the functional connectivity between the brain activations S for subject s and the average brain activations of all other subjects O using TimeCorr ISFC

with variance σ. The correlation between S_t^i—activation of brain region i for subject s at time t—and O_t^j—average brain activation of brain region j of other subjects at time points t—is obtained through the following equation:

$$C(S_t^i, O_t^j) = \frac{1}{Z} \frac{\sum_{l=0}^{T} (S_l^i - \bar{S_t^i}) \cdot (O_l^j - \bar{O_t^j}) \cdot \mathcal{N}(l|t, \sigma)}{\sigma_{S_t^i} \cdot \sigma_{O_t^j}}$$

Where

$$Z = \sum_{l=0}^{T} \mathcal{N}(l|t, \sigma)$$

$$\bar{S_t^i} = \frac{1}{Z} \sum_{l=0}^{T} S_l^i \cdot \mathcal{N}(l|t, \sigma)$$

$$\bar{O_t^i} = \frac{1}{Z} \sum_{l=0}^{T} O_l^i \cdot \mathcal{N}(l|t, \sigma)$$

$$\sigma_{S_t^i} = \sqrt{\frac{1}{Z} \sum_{l=0}^{T} (S_l^i - \bar{S_t^i})^2 \cdot \mathcal{N}(l|t, \sigma)}$$

$$\sigma_{O_t^i} = \sqrt{\frac{1}{Z} \sum_{l=0}^{T} (O_l^i - \bar{O_t^i})^2 \cdot \mathcal{N}(l|t, \sigma)}$$

(c) Repeat the above process for every pair of brain regions, for every time point.

3. Apply Fisher Z-transformation to every element r of the correlation matrices for each subject at each time points to obtain the corresponding Z-correlation matrices:

$$z = \frac{1}{2} \ln(\frac{1+r}{1-r})$$

4. Average the Z-correlation matrices Z_i across all subjects:

$$S_Z = \frac{1}{N} \sum_{i=1}^{N} Z_i$$

5. Apply inverse Z-transformation to the average Z-correlation matrix to obtain the Inter-subject Functional Connectivity (ISFC) mean correlation matrix:

$$ISFC = \frac{\exp(S_Z + S_Z^T) - 1}{\exp(S_Z + S_Z^T) + 1}$$

6. Convert the correlation matrix at every time point to its inverse squareform format, and output a T by $(V^2 - V)/2$ dimensional matrix containing the average functional connectivity of all subjects in the dataset, where T and V represent the time length and brain regions count of the dataset, respectively.

2.3 Multi-Subject TimeCorr Level-Up

The multi-subject TimeCorr Level-Up process utilizes the functionalities of the single-subject TimeCorr method to extract high-order brain dynamics patterns from brain fMRI data. Due to the data-losing flaw of the sliding window method, research related to functional connectivity heretofore have been largely limited to analysis on raw fMRI activations and first level functional connectivity. However, with the availability of TimeCorr, we are able to repeatedly find higher order functional connectivity of an fMRI dataset without fear of losing data with each "level up". This advantage allows us to explore higher-order dynamic patterns within the brain that was previously impossible to access.

The concept of "leveling up" involves repeatedly finding the correlation of correlations of brain activity. Given a brain fRMI dataset, we designate the raw activations as level 0 and the functional connectivity (correlation structure) of the activations as level 1, the result of "leveling up" from level 0 (here "level" and "order" are interchangeable). Hereafter, given a matrix of functional connectivity of level n, we can apply the TimeCorr Level-Up method to find the correlation structure of the data at level n to obtain functional connectivity at level $n + 1$.

In addition to effectively retaining all relevant data, the Multi-Subject TimeCorr Level-Up method also presents a scalable solution to large datasets. Normally, for an input dataset of V voxels (brain regions), the first order functional connectivity would occupy V^2 space and every additional level of functional connectivity would involve an exponential increase in storage requirement. In order to circumvent this problem, the Multi-Subject TimeCorr Level-Up method first finds the inverse square-form (vectorized array containing the non-diagonal elements of a square matrix) of the correlation matrices for every time point in the "level up" output, and then uses PCA to further reduce the result to a matrix of T by V dimensions, where T and V represent the time length and voxel count of the input dataset.

This solution ensures all orders of functional connectivity derived from the same dataset will have the same dimensions. In addition, the application of PCA further guarantees that only the most principle components of each functional connectivity matrix is retained for further analysis.

Formal Definition

Given a brain activation or activation correlation matrix at level l of dimensions S subject, V brain regions and T time points, the Level-Up function

1. Applies single-subject TimeCorr on the data matrix for each subject to obtain correlations matrix
2. Concatenate the correlations matrix for each subject from the previous step together along the first dimension into a single matrix of dimensions $(S * V)xT$
3. Apply PCA on the concatenated matrix from the previous step to obtain a reduced representation of the correlation matrix
4. Separate the reduced correlation matrix from the previous step into data for each subject to obtain a 3-D matrix of size SxVxT
5. Repeat the above process on the output to level up again

2.4 Multi-Subject Decoding Analysis with TimeCorr ISFC

The Multi-Subject Decoding Analysis is a method designed to find stimulus-dependent activations in fMRI dataset by cross referencing data from multiple subjects [1]. Through application of Multi-Subject TimeCorr ISFC on two equal divisions of the subject data, "white noises" (background noise, human error, random neural activity, etc) are dampened and activities that are common across subjects are emphasized, thereby increasing the salience of patterns that reflect activities that are only stimulus-driven. Additionally, when the correlation between the same time point in both sub-groups is high, it indicates that the average brain activities in one group demonstrates very high similarity with the average brain activities of all subjects in the other group at that time point. When the

brain activities of multiple subjects display high similarity for certain time points in their brain activations time sequence, then intuitively there's a high probability that their brain responses for those time points are uniformly caused by their common stimulus. Using this logic, the decoding accuracy of the dataset, percentage of time points that demonstrates highest correlation with themselves across two sub-division of subject data, gives an accurate representation of

1. Similarity between subject responses throughout the fMRI simulation
2. The average proportion of subject activities within the dataset that are stimulus dependent
3. The potency of subjects' brain response to the stimuli
4. The average proportion of subject activities that shows significant distinction from the rest of the time sequence
5. The potency of stimuli in generating salient cognitive response across all subjects

Due to the above characteristics, the Decoding Analysis makes up a crucial part of our project as a means to evaluate the effectiveness of our newly developed methods.

Formal Definition

Given a multi-subject dataset containing S subjects, each possessing T time points and V voxels, to conduct Decoding Analysis:

1. Select a Gaussian density variance V for TimeCorr coefficient generation.
2. For n repetitions:
   (a) Randomly divide the subjects into two equal groups.
   (b) Calculate Inter-Subject Functional Connectivity (ISFC) for each group using TimeCorr. The resulting ISFC matrices are labelled I_1 and I_2.
   (c) Calculate time-point-wise correlation between functional connectivity at each time point in I_1 and functional connectivity at each time point in I_2, resulting in a T by T correlation matrix.

(d) A time point in I_1 is correctly decoded if the time point of highest correlation in I_2 is the same time point, and vice versa for I_2.

(e) Count and record the number of time points that were correctly decoded for I_1 and I_2.

3. Sum the total number of time points correctly decoded for I_1 and I_2 across all repetitions, and find the average through dividing by the product between the repetition count and two times the time point count (to account for both I_1 and I_2). The result is the decoding accuracy output of the multi-subject dataset.

3 Results

3.1 Testing Single Subject TimeCorr on Synthetic Datasets

To conduct a side-by-side comparison of the correlation recovery functionalities between the TimeCorr method and the traditional sliding window method, we constructed synthetic datasets with pre-defined correlation patterns using Cholesky decomposition. This approach takes an original X variable (or matrix) and uses the Cholesky transformation to create a new, correlated, Y variable. Using this approach, we are able generate a time sequence of synthetic brain activations that exactly follows a pre-designated a dynamically correlation structure.

Two kinds of synthetic datasets were generated to test different features of TimeCorr: datasets with multiple time blocks of different correlations and datasets with ramping correlation structures.

3.1.1 Single Subject Synthetic Dataset with block correlations

In this dataset, we divide time sequence into N blocks of equal time lengths. Each block possesses a different correlation that's consistent throughout the length of the block. This dataset will be used to test TimeCorr's ability to distinguish different correlations within a dataset. To generate the dataset:

1. Using random normal distribution, generate a dataset X of dimensions T by V, where T and V represents the desired time length and voxel (brain region) count, respectively.
2. Design a correlation matrix of dimensions V by V for each of the N blocks, denoted R_n.
3. For each of the N blocks in X, transform the block X_n into correlated structure through the equation:

$$W_n = Cholesky(R_n) \cdot X_n$$

Where $Cholesky(R_n)$ represents the Cholesky decomposition of the correlation matrix R_n

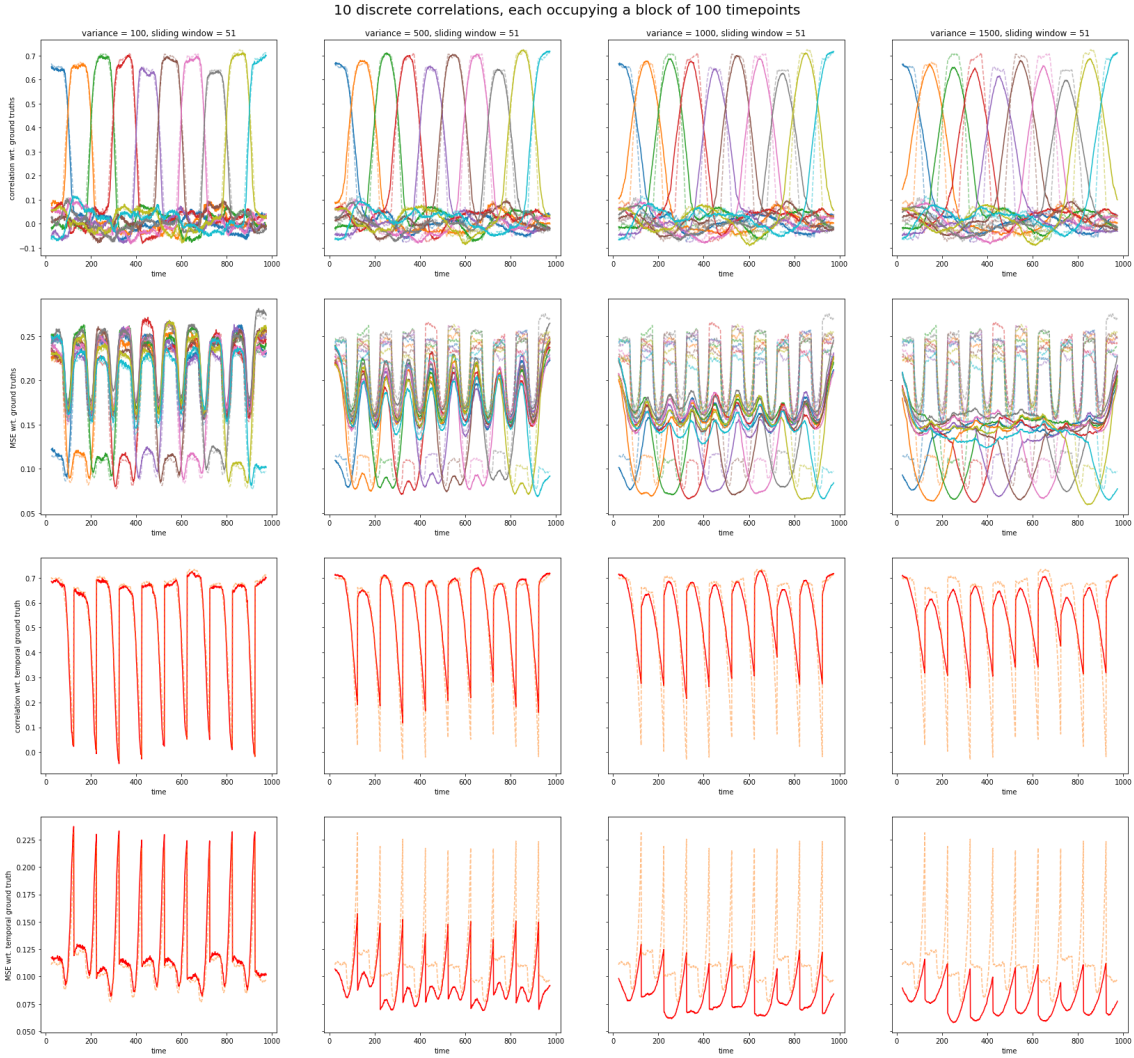
4. Piecing all of the blocks together to form the new synthetic dataset W that possesses the designated dynamic correlation structure.

The average results of running TimeCorr and the sliding window method on 100 datasets containing 10 blocks, each occupying 100 time points, is displayed in Figure 1. In the graphs on the first row, each color represents the correlation between the correlation structure of a block with the solutions from the TimeCorr and the sliding window approach, which are represented by solid and dotted lines, respectively. The graphs show that both the TimeCorr and the sliding window approach are able to effectively recover the correct correlation structure for each block. In addition, the TimeCorr approach is able to more accurately recover the temporal block correlation at each time point when the Gaussian density variance for its coefficient generation is low (in other words, when the resolution is high). In contrast, when the TimeCorr Gaussian variance increases, the solutions slowly gains smoother transition in-between correlation blocks but in exchage for lower recovery resolution and lower temporal accuracy.

These results are confirmed by the results shown in the second row, which shows the MSE between the block correlations and the TimeCorr and the sliding window solutions, which are represented by solid and dotted lines, respectively. TimeCorr and the sliding window approach seem to have very similar performances when the TimeCorr variance is low. But as the Guassian density variance for TimeCorr coefficients generation increases, the mean squared error of TimeCorr decreases significantly, indicating that TimeCorr is able to produce more precise solutions at block centers when the Gaussian density variance is high.

The graphs in the third and fourth row show the correlation and MSE between the temporal ground truth (true correlation at each time point) and the TimeCorr and the sliding window solutions, represented by solid and dotted lines respectively. As the Gaussian variance of TimeCorr increases, the correlation between its solution and the ground truth at transitions improves drastically, and the mean squared error experiences significant improvements across the board. When comparing the TimeCorr results with the sliding window results in these two rows, we can see that TimeCorr performs significantly better in both correlation

Figure 1: Testing on 10 blocks of different correlations

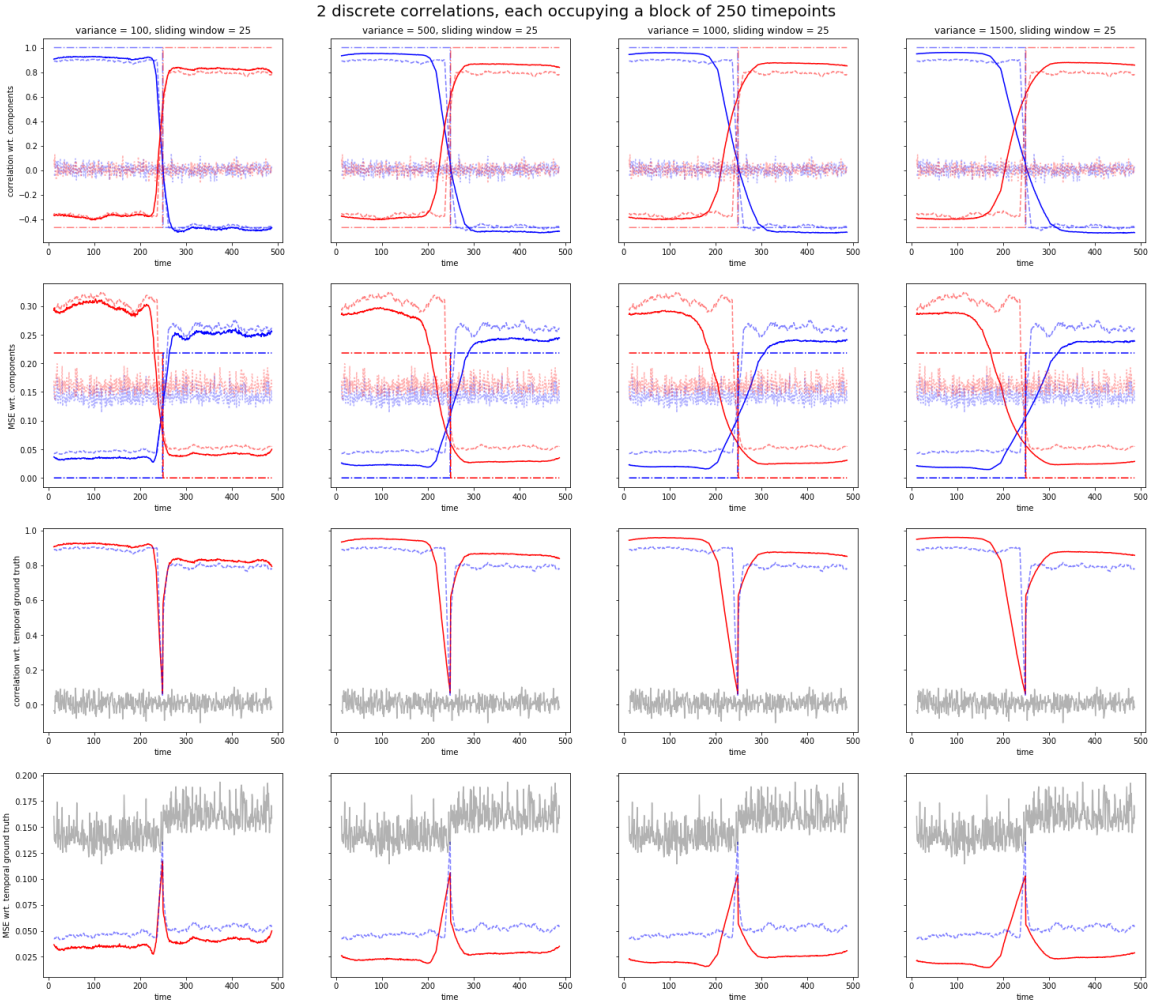


(a) In the first row, solid lines show the correlation between block correlations and TimeCorr solutions; dotted, sliding window solutions. (b) In the second row, solid lines show mean squared error(MSE) between block correlations and TimeCorr solutions; dotted, sliding window solutions. (c) In the third row, the solid line shows correlation between temporal ground truth and TimeCorr solutions; dotted, sliding window solutions. (d) In the fourth row, the solid line shows MSE between temporal ground truth and TimeCorr solutions; dotted line, sliding window solutions. (e) The columns represent TimeCorr implementations with a different Gaussian density variances.

and MSE when the TimeCorr variance is high.

To further confirm our results, we rerun the test on 100 dataset containing 2 blocks of

Figure 2: Testing on 2 blocks of different correlations



(a) In the first row, solid lines show the correlation between block correlations and TimeCorr solutions; dotted, sliding window solutions; dot-dashed, ground truth. (b) In the second row, solid lines show mean squared error(MSE) between block correlations and TimeCorr solutions; dotted, sliding window solutions; dot-dashed, ground truth. (c) In the third row, the solid line shows correlation between temporal ground truth and TimeCorr solutions; dotted, sliding window solutions; gray, random guess. (d) In the fourth row, the solid line shows MSE between temporal ground truth and TimeCorr solutions; dotted line, sliding window solutions; gray, random guess. (e) The columns represent TimeCorr implementations with a different Gaussian density variances.

different correlations, each containing 250 time points. The average results are shown in Figure 2, with ground truth and random guess displayed for comparison. From these

graphs, we can can confirm TimeCorr boasts significantly better performances than the sliding window approach in both correlation and MSE when the ground truth correlation is stable. Furthermore, TimeCorr is able to more effectively recover correlations structures at abrupt transitions using a high resolution setup (low Gaussian variance).

3.1.2 Single Subject Synthetic Dataset with ramping correlations

In this dataset, we adopt a correlation structure that linearly changes from one correlation C_1 to another correlation C_2. This dataset will be used to test TimeCorr's ability to accurately recover correlation structures that are dynamically changing over time.

1. Using random normal distribution, generate a dataset X of dimensions T by V, where T and V represents the desired time length and voxel (brain region) count, respectively.

2. Generate two correlation matrices C_1 and C_2, then generate temporal correlation structures using the function

$$C_t = z2r(\frac{(T-t)\cdot r2z(C_1)}{T} + \frac{t\cdot r2z(C_2)}{T})$$

   where T is the total time length of the datase. $r2z$ and $z2r$ represent Fisher Z-transformation and inverse Fisher Z-transformation, respectively. These transformations were applied to ensure stable combination of correlation structures.

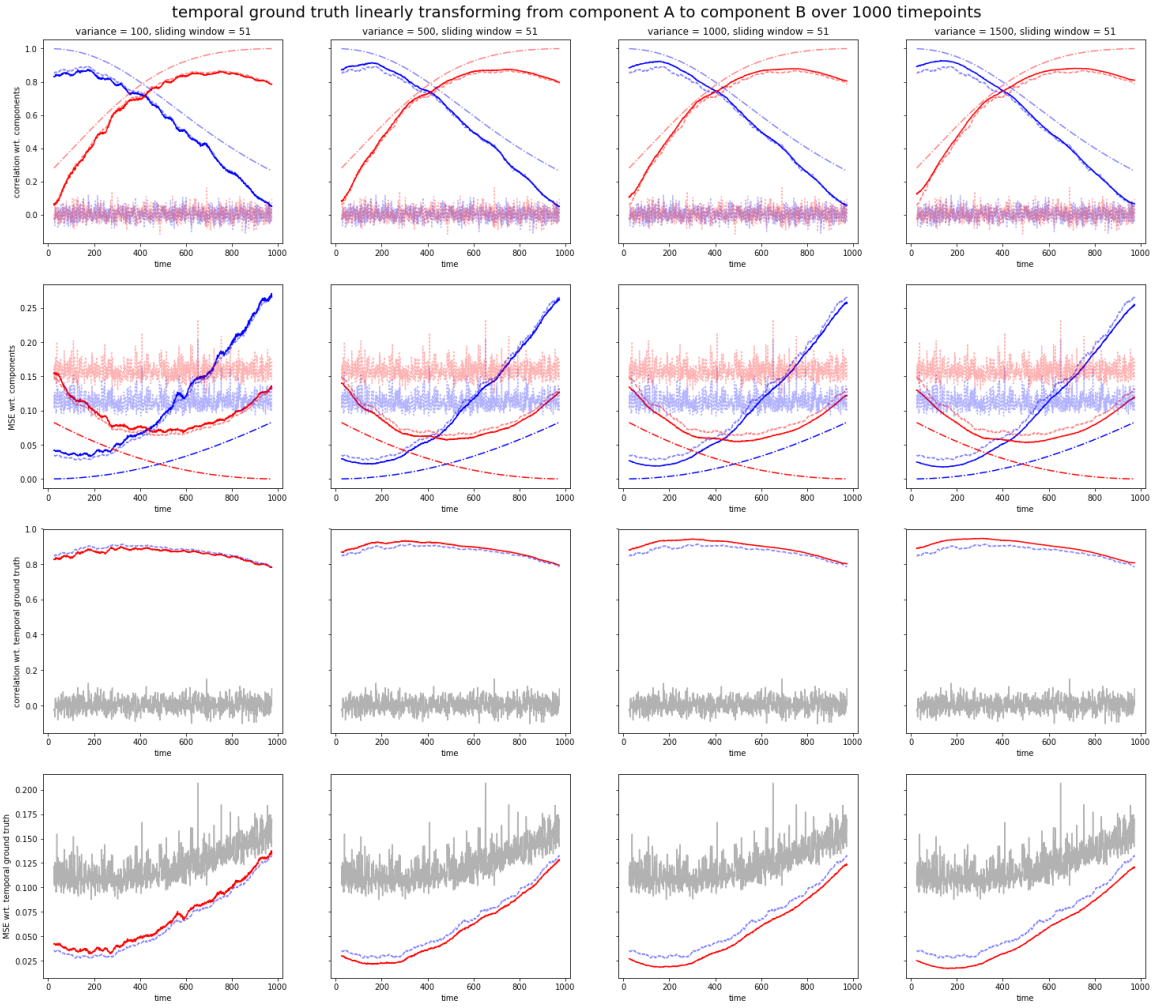
3. For activations X_t at each time point in X, transform the data into correlated structure through the process:

$$W_t = Cholesky(C_t)\cdot X_t$$

4. Piecing all of the blocks together to form the new synthetic dataset W that possesses the designated dynamic correlation structure.

First, 100 ramping datasets with 1000 time points were generated to produce a gradual linear transition between two distinct correlation structures. The average results are displayed in Figure 3. At Gaussian variance equal to 100, TimeCorr shows very similar performance

Figure 3: Testing on 1000 time point ramping dataset



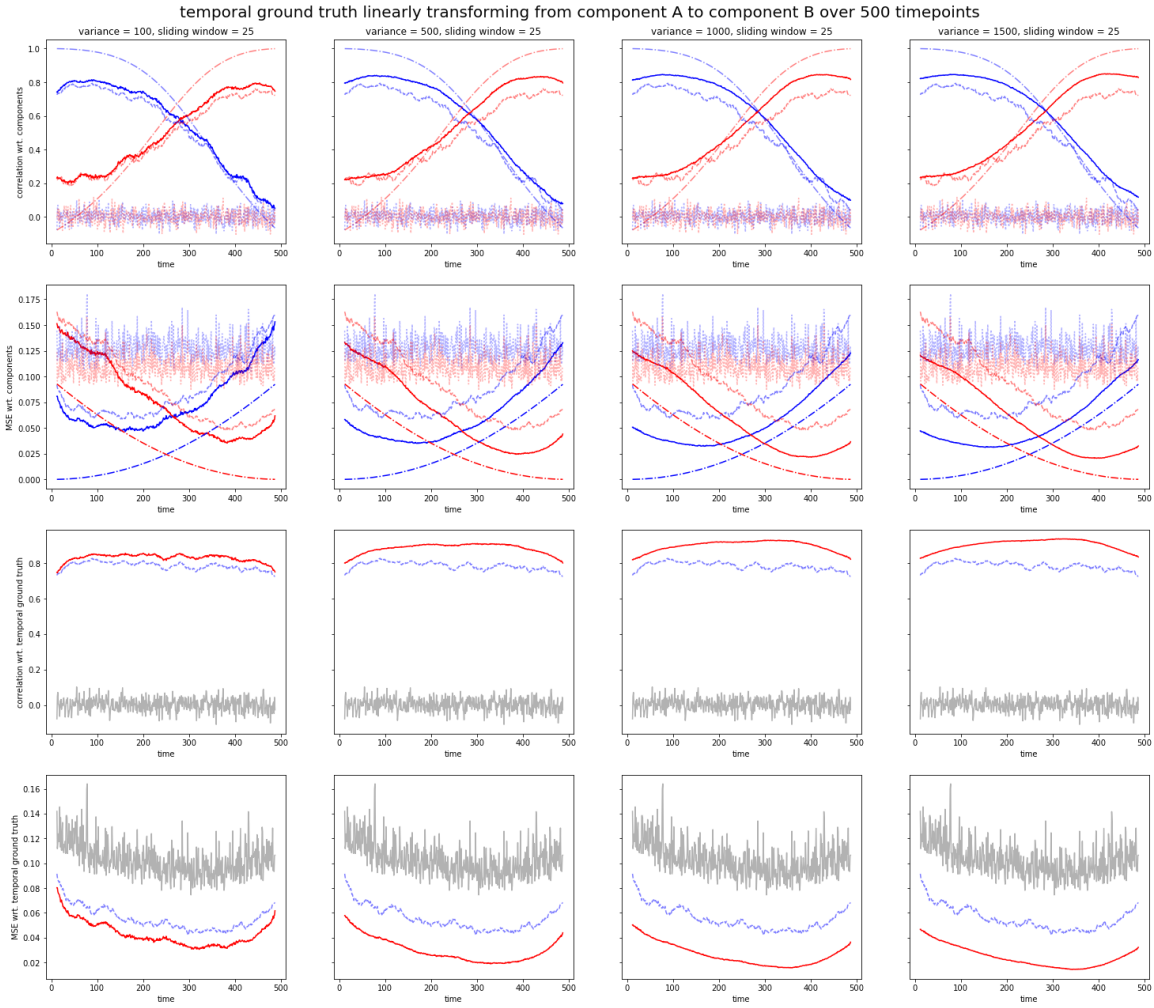
(a) In the first row, solid lines show the correlation between block correlations and TimeCorr solutions; dotted, sliding window solutions; dot-dashed, ground truth. (b) In the second row, solid lines show mean squared error(MSE) between block correlations and TimeCorr solutions; dotted, sliding window solutions; dot-dashed, ground truth. (c) In the third row, the solid line shows correlation between temporal ground truth and TimeCorr solutions; dotted, sliding window solutions; gray, random guess. (d) In the fourth row, the solid line shows MSE between temporal ground truth and TimeCorr solutions; dotted line, sliding window solutions; gray, random guess. (e) The columns represent TimeCorr implementations with a different Gaussian density variances. (f) In the first two rows, the blue lines represent the relationship between recovered solution and the left terminal correlation; red line, right terminal correlation.

with the sliding window method with a window length of 51, albeit TimeCorr results do not come at the expense of extensive data-loss. When the Gaussian variance of TimeCorr is increased to above 100, the TimeCorr solutions' advantages over the sliding window solution immediately becomes obvious in both MSE and their correlation with the temporal ground truths. This phenomenon is most evident in the third and fourth row, where the TimeCorr solutions consistently shows higher correlation and lower MSE with the ground truth at almost every time point.

TimeCorr's performance advantage is further verified by the testing results from 100 ramping datasets of 500 time points, where the linear transitions between the two distinct terminal correlation structures are accelerated from the 1000 time point datasets. The average results are shown in Figure 4. In contrast to the results from the 1000 time point datasets, the performance difference is more evident in the results from the 500 time point datasets. As the temporal ground truth correlation linearly transitions from the left terminal correlation structure to the right terminal correlation structure within 500 time points, the rate of change at each time point is much greater than that of the 1000 time point dataset. Under these conditions, the TimeCorr results consistently shows significantly higher correlation and lower MSE with the temporal ground truths than the sliding window results. This result indicates that the TimeCorr approach is better at recovering rapidly changing dynamic correlation structures than the sliding window approach. In addition, the results does not seem to change significantly from the 500 variance TimeCorr implementation to the 1500 variance TimeCorr implementation, which might be due to the variance exceeding the dataset time length.

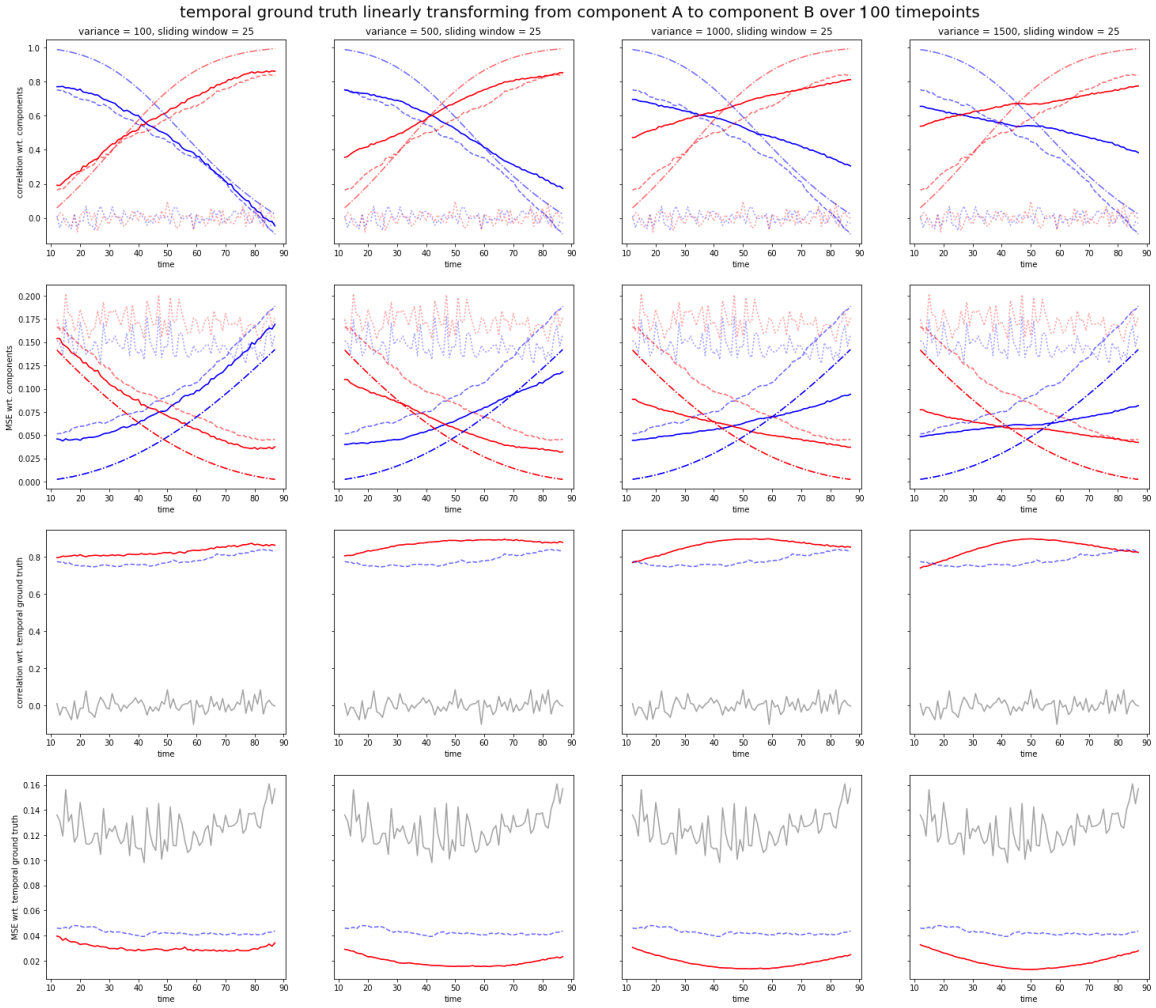
To further verify the advantages of TimeCorr in recovering highly dynamic correlation structures and the limitations of changing TimeCorr Gaussian variance in datasets with short time lengths, we conducted analysis on 100 ramping datasets of only 100 time points, where the linear transitions between the two distinct terminal correlation structures are further accelerated from the 500 time point datasets. The results are shown in Figure 5 and confirm our observations from the previous tests: the TimeCorr approach consistently out-performs the sliding window approach in every setup and TimeCorr's performance stagnates/worsens as its Gaussian variance increases beyond the total time length of the dataset. These results further consolidates TimeCorr's superiority over the sliding window method in recovering

Figure 4: Testing on 500 time point ramping dataset



(a) In the first row, solid lines show the correlation between block correlations and TimeCorr solutions; dotted, sliding window solutions; dot-dashed, ground truth. (b) In the second row, solid lines show mean squared error(MSE) between block correlations and TimeCorr solutions; dotted, sliding window solutions; dot-dashed, ground truth. (c) In the third row, the solid line shows correlation between temporal ground truth and TimeCorr solutions; dotted, sliding window solutions; gray, random guess. (d) In the fourth row, the solid line shows MSE between temporal ground truth and TimeCorr solutions; dotted line, sliding window solutions; gray, random guess. (e) The columns represent TimeCorr implementations with a different Gaussian density variances. (f) In the first two rows, the blue lines represent the relationship between recovered solution and the left terminal correlation; red line, right terminal correlation.

Figure 5: Testing on 100 time point ramping dataset



(a) In the first row, solid lines show the correlation between block correlations and TimeCorr solutions; dotted, sliding window solutions; dot-dashed, ground truth. (b) In the second row, solid lines show mean squared error(MSE) between block correlations and TimeCorr solutions; dotted, sliding window solutions; dot-dashed, ground truth. (c) In the third row, the solid line shows correlation between temporal ground truth and TimeCorr solutions; dotted, sliding window solutions; gray, random guess. (d) In the fourth row, the solid line shows MSE between temporal ground truth and TimeCorr solutions; dotted line, sliding window solutions; gray, random guess. (e) The columns represent TimeCorr implementations with a different Gaussian density variances. (f) In the first two rows, the blue lines represent the relationship between recovered solution and the left terminal correlation; red line, right terminal correlation.

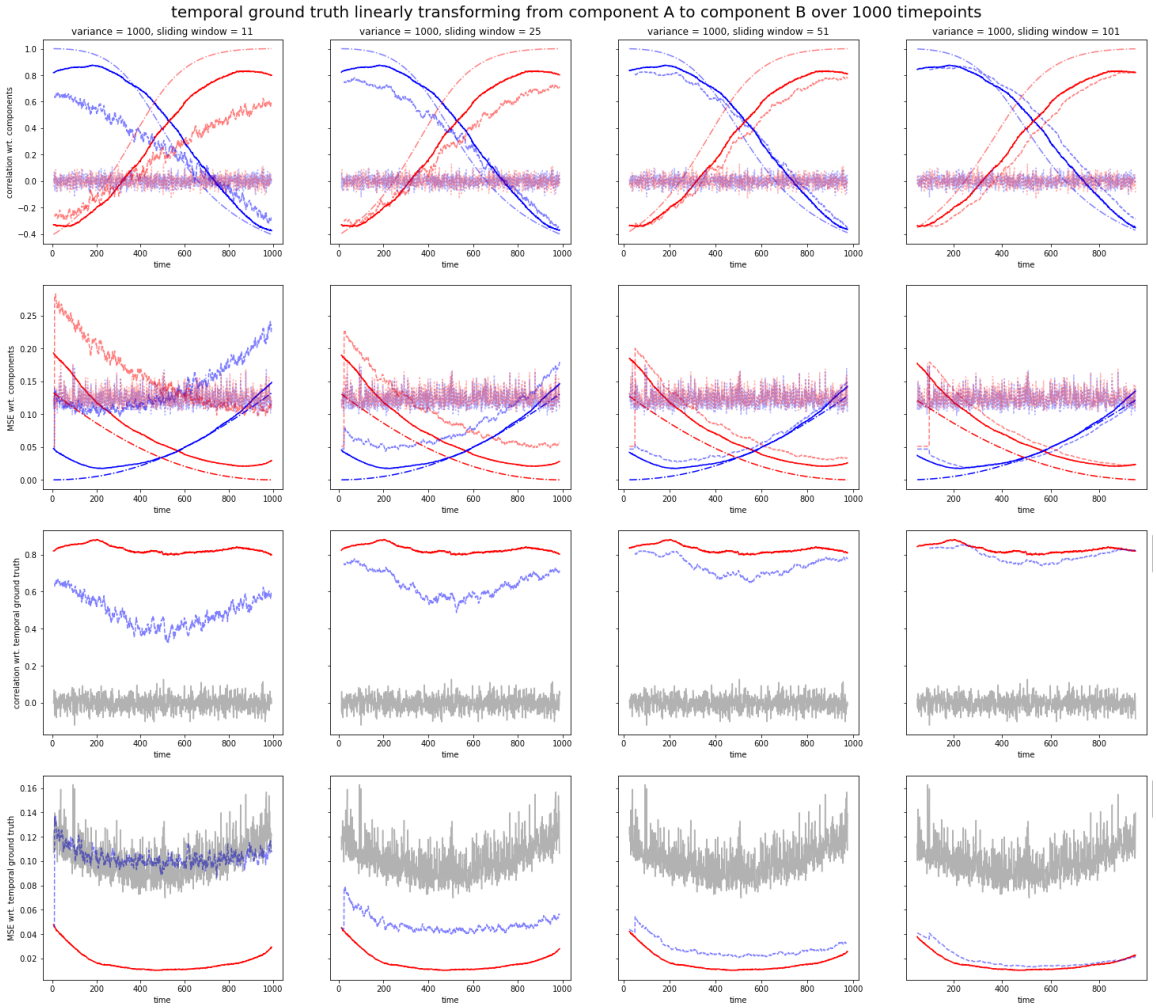
highly dynamic correlation structures. In addition, we gained very useful insight on how to select Gaussian variance for TimeCorr to produce the best recovery results: the most stable results seem to come from a Gaussian variance equal to the lesser between the total time length and 1000 time points.

Lastly, to demonstrate TimeCorr's advantage over the sliding window approach generalizes to all setups, we conducted an analysis on 100 ramping datasets, each containing 1000 time points, to compare TimeCorr with sliding window implementations of varying window length. The results are shown in Figure 6. The TimeCorr setup used for comparison has a Gaussian variance of 1000, a standard setup with average performance among all TimeCorr setups in the previous tests. From the graphs, we can see that the performance of sliding window generally increases as the window length increases, but the rate of improvement decreases dramatically past a window length of 51 time points (marginal improvement when an additional 50 time points is added to the window length). In addition, the performance of the sliding window approach at a window length of 11 time points is underwhelming in comparison with the TimeCorr approach. As the window length increases to 101 time points, the correlation and MSE between the sliding window solution and temporal ground truths gradually nears the performance of the standard TimeCorr implementation. However, a window length of 101 time points is equivalent to around 10% of the time length of most fMRI datasets. Losing 10% of the information for calculation of functional connectivity at each level is impractical for explorations of high level brain dynamics. For the above reasons, we chose TimeCorr as the more preferable method of functional connectivity calculation for our High-Order Brain Dynamics Model.

3.2 Testing Inter-Subject Functional Connectivity using TimeCorr

Confident with the TimeCorr approach's ability to recover the dynamic correlation structures of single subject datasets, we proceeded to test TimeCorr's performance in recovering inter-subject functional connectivity in multi-subject synthetic datasets. In practical scenarios, fMRI datasets always contain a significant amount of noise from the environment, human error, random neural activity, etc in addition to stimulus-driven brain activities. Therefore, to simulate realistic human brain activities, we added different levels of random

Figure 6: Testing on 100 time point ramping dataset



(a) In the first row, solid lines show the correlation between block correlations and TimeCorr solutions; dotted, sliding window solutions; dot-dashed, ground truth. (b) In the second row, solid lines show mean squared error(MSE) between block correlations and TimeCorr solutions; dotted, sliding window solutions; dot-dashed, ground truth. (c) In the third row, the solid line shows correlation between temporal ground truth and TimeCorr solutions; dotted, sliding window solutions; gray, random guess. (d) In the fourth row, the solid line shows MSE between temporal ground truth and TimeCorr solutions; dotted line, sliding window solutions; gray, random guess. (e) The columns represent different sliding window implementations with varying window lengths. (f) In the first two rows, the blue lines represent the relationship between recovered solution and the left terminal correlation; red line, right terminal correlation.

noise to the synthetic datasets to gauge the robustness of TimeCorr ISFC in recovering stimulus-related functional connectivity. The noise levels we experimented with were on the order of magnitudes of 10% (0.1), 100%(1), 1000%(10) and 10000%(100) of the average activation magnitude. Running TimeCorr with Gaussian variance of 1000 time points and a sliding window setup with a window length of 25 time point on 100 ramping datasets, each containing 300 time points, we obtained the results displayed in Figure 7.

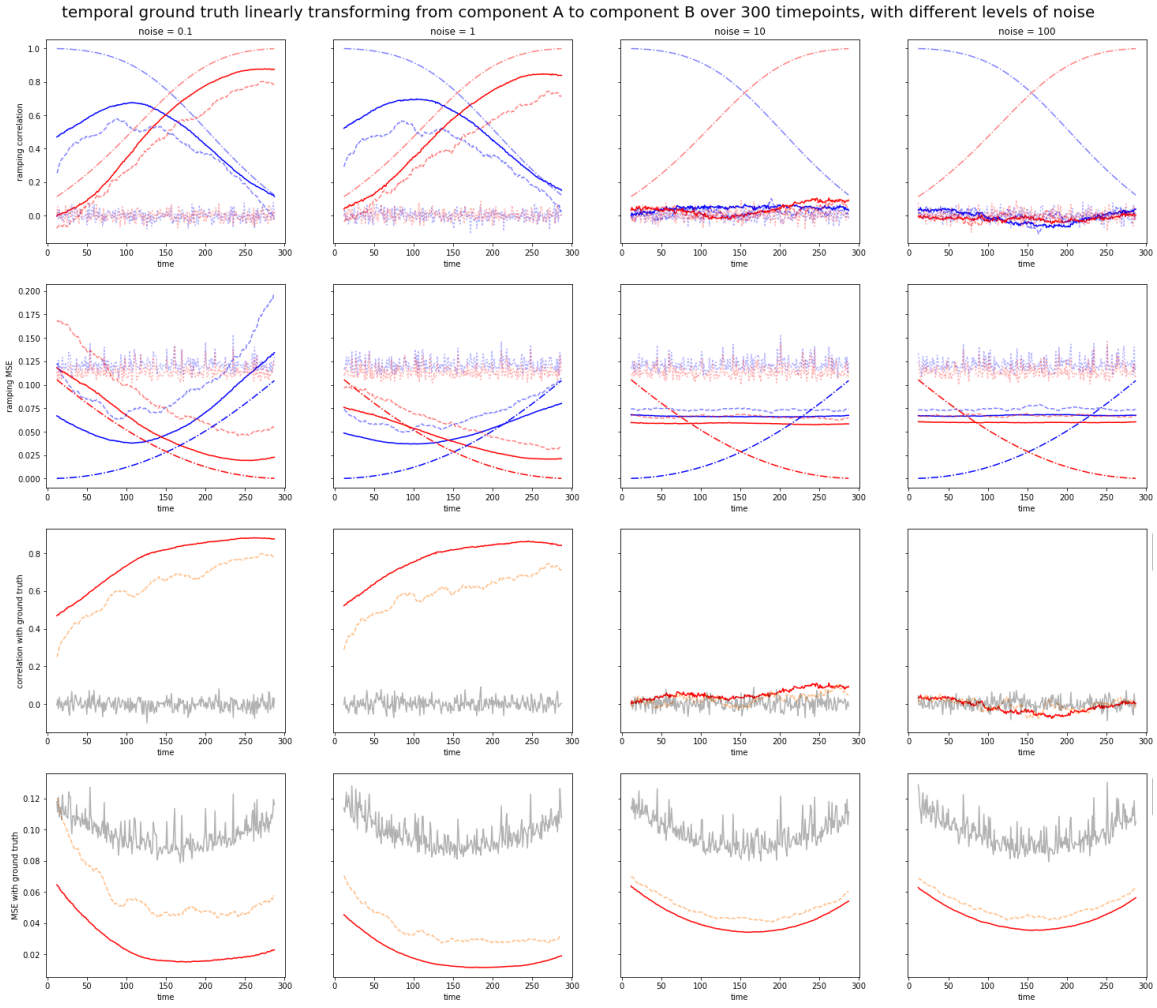
The results show that TimeCorr is able to retrieve the stimulus-driven dynamic correlation structure with relatively high accuracy at low(0.1) to medium(1) noise levels, but fails to recover any meaningful information when the noise level is too high(10 and 100). These results highlight the difficulty of recovering stimulus-driven correlation structures when the subjects display weak cognitive response to the stimulus. However, when the stimuli is able to evoke reasonably strong cognitive response from the subject, TimeCorr can effectively recover the stimuli-driven dynamic correlation structures with relatively high accuracy.

In addition, due to the limited number of total time points in the dataset, we chose to use a sliding window setup with a window length of 25 time points. When comparing the correlation and MSE between their solutions and the temporal ground truths, the TimeCorr approach distinctly outperforms the sliding window method in every situation. Furthermore, as the amount of noise increases, the performance of the sliding window approach deteriorates more noticeably than that of the TimeCorr approach. These contrasts further consolidates TimeCorr as the superior choice for functional connectivity calculations.

3.3 Decoding Analysis using TimeCorr ISFC

Next, we evaluated how brain dynamics is represented by each level of high-order functional connectivity. On top of the PCA reduced activations from fMRI datasets, we used the TimeCorr Level Up function to generate 10 levels of high-order functional connectivity, each containing an equal number of voxels. Next, we conducted decoding analysis using a TimeCorr setup of on each level of information, from PCA reduced fMRI activations (level 0) to the 10th order functional connectivity (level 11). Three TimeCorr setups were implemented—each with Gaussian variances of 1000 time points, 10 time points, or 1 time point—to comprehensively understand the decoding capability of each level under varying

Figure 7: Testing on 100 time point ramping dataset



(a) In the first row, solid lines show the correlation between block correlations and TimeCorr solutions; dotted, sliding window solutions; dot-dashed, ground truth. (b) In the second row, solid lines show mean squared error(MSE) between block correlations and TimeCorr solutions; dotted, sliding window solutions; dot-dashed, ground truth. (c) In the third row, the solid line shows correlation between temporal ground truth and TimeCorr solutions; dotted, sliding window solutions; gray, random guess. (d) In the fourth row, the solid line shows MSE between temporal ground truth and TimeCorr solutions; dotted line, sliding window solutions; gray, random guess. (e) The columns represent recovery results from datasets with varying noise levels. (f) In the first two rows, the blue lines represent the relationship between recovered solution and the left terminal correlation; red line, right terminal correlation.

resolutions.

Three different datasets were used for this analysis: Sherlock, Pieman, and Forrest. Detailed descriptions for each dataset are included in each subsection. For each TimeCorr setup for each dataset, we conducted 100 repetitions of decoding analysis with different random group divisions. All of the following results are the average of 100 repetitions.

3.3.1 Description for the Pieman Dataset

The stimuli for this experiment were derived from a 7 minute recording of the "Pie Man" story by Jim O'Grady. The experiment was conducted on native English speakers with normal hearing, and was divided into four testing conditions: 36 subjects listened to the story from beginning to end, and were labelled as the "intact" group; 18 subjects listened to a version of the story where the paragraphs were scrambled randomly, and were labelled as the "paragraph" group; 36 subjects listened to a version of the story where the words were scrambled randomly, and were labelled as the "word" group; lastly, 36 subjects did not listen to anything and stayed in resting state throughout the experiment, and were labelled as the "resting" group. 12 seconds of neutral music and 3 seconds of silence preceded and 15 seconds of silence followed each playback in all conditions. The data from music and silence periods were discarded from all analyses.[5]

3.3.2 Description for the Forrest Dataset

The stimuli for this experiment is a two-hour presentation of an audio-only version of the Hollywood feature film "Forrest Gump" made specifically for a visually impaired audience. The movie was divided into eight 15-minute sessions with breaks in-between, and presented to 20 subjects with normal hearing. During each session, brain activities centered on the auditory cortices in both brain hemisphere and the frontal and posterior portions of the brain were recorded into high resolution 7-Tesla fMRI datasets.[33]

3.3.3 Description for the Sherlock Dataset

Sixteen participants (17 were described in the original paper, but 1 was removed for this dataset due to a small amount of missing data at the end of the movie scan.) were presented with a 50-minute segment of the audio-visual movie Sherlock (BBC) while undergoing fMRI scan. Before each viewing, the subjects were informed that they would later need to describe the contents of movie. Following the viewing, participants were asked to describe aloud what they recalled of the movie in as much detail as they could while undergoing brain imaging, with no visual input or experimenter guidance. Participants were allowed to speak on any aspect of the movie for as long as they wished, while their speech was recorded with an fMRI-compatible microphone. [32]

3.3.4 Results

1. The average results for each TimeCorr setup implemented on the Pieman dataset are displayed in Figure 8, 9, and 10.

2. The average results for each TimeCorr setup implemented on the Forrest dataset are displayed in Figure 11, Figure 12 and Figure 13.

3. The average results for each TimeCorr setup implemented on the Sherlock dataset are displayed in Figure 18, Figure 19 and Figure 14.

When comparing our results with previous efforts, we can see that the decoding accuracy we achieved on the first order ISFC for the Pieman dataset using TimeCorr-based methods is significantly higher across all TimeCorr setups [1]. As higher decoding accuracy is a strong indication of the richness of stimulus-driven activities within a dataset, our improved results further confirm that TimeCorr-based methods are superior in the recovery of stimulus-driven correlation structure in fMRI datasets. Furthermore, the decoding accuracy we achieved on PCA reduced fMRI data is significantly higher than the decoding accuracy obtained from HTFA reduced datasets, indicating that PCA is more effective in preserving dynamic structure than HTFA and justifying it's usage in our level up method.

Figure 8: Decoding Analysis using TimeCorr with variance of 1000 time points

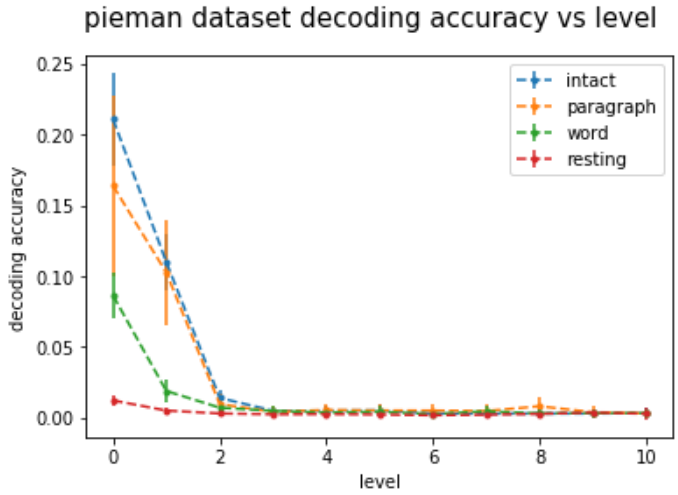
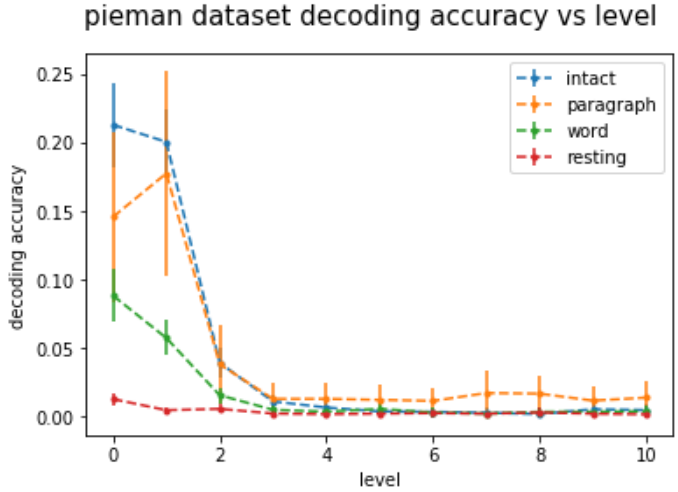


Figure 9: Decoding Analysis using TimeCorr with variance of 10 time points



Another important observation that is consistent across results from all three datasets is that the higher order functional connectivities actually produced lower decoding accuracy. This result was unexpected as brain dynamics across subjects should intuitively become more similar at higher orders. One possible explanation for this is, although the high order functional connectivity still contain information about stimulus-driven brain activities, there seems to have been a small amount of information loss in each level up process, leading to slightly lower classification accuracy. However, although higher order functional connectivities display weaker decoding capabilities than lower level functional connectivities, our hypothesis is that each order of functional connectivity contains distinct information about

Figure 10: Decoding Analysis using TimeCorr with variance of 1 time point

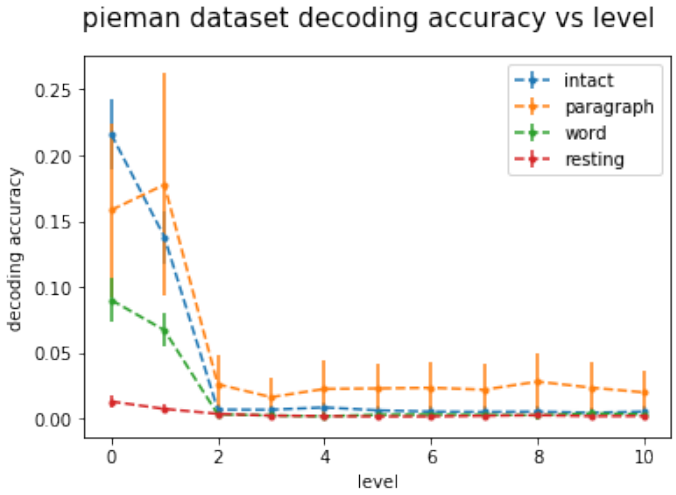
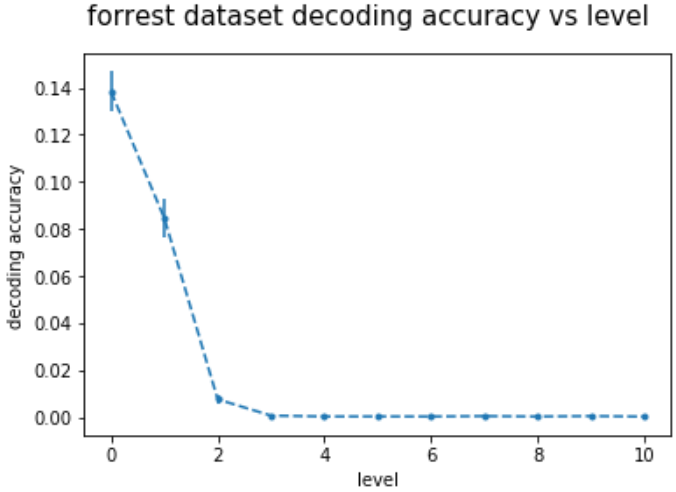


Figure 11: Decoding Analysis using TimeCorr with variance of 1000 time points



brain dynamics. If this is the case, then a mixture of all orders of functional connectivity will potentially have stronger decoding performance than that of functional connectivity at any single level.

We also suspected that TimeCorr temporal resolution had some influence on the results, which is why our analysis of each dataset consisted of a variety of different TimeCorr setups. Our hypothesis is that there is a loss in information in each level up process, therefore the loss would intuitively grow exponentially as we reach higher orders of functional connectivity. Therefore, TimeCorr with Gaussian variances of 1000 time points, 10 time points and 1 time point were chosen to explore if variation in TimeCorr temporal resolution would affect

Figure 12: Decoding Analysis using TimeCorr with variance of 10 time points

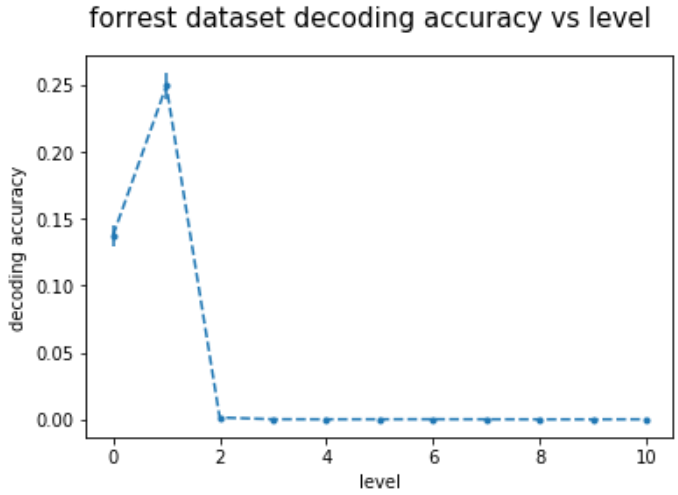
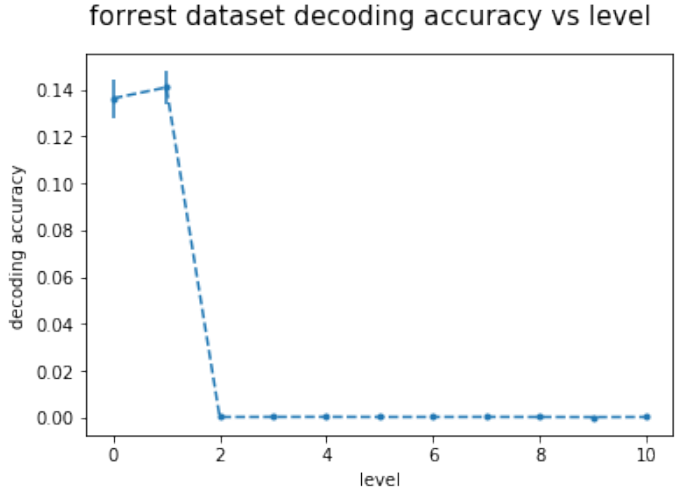


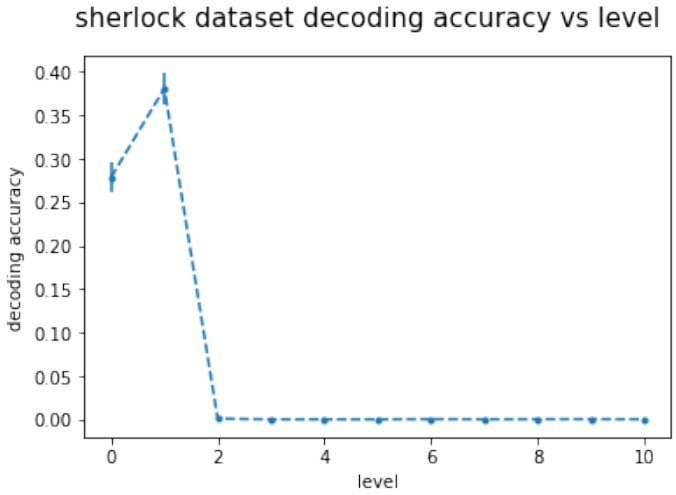
Figure 13: Decoding Analysis using TimeCorr with variance of 1 time point



the distribution of decoding accuracy across different levels.

The results from the Pieman dataset showed that the decoding accuracy of lower orders of functional connectivity increasing drastically when the TimeCorr temporal resolution increased from a Gaussian variance of 1000 time points to a Gaussian variance of 10 time points, but undergoing a decrease as the temporal resolution was increased further to a Gaussian variance of 1 time point. The results from Sherlock and Forrest datasets also show similar patterns. Additionally, we observe that as the TimeCorr temporal resolution increases from Gaussian variance of 1000 time points to Gaussian variance of 1 time point, the decoding accuracy of higher order functional connectivities of the Pieman dataset seem

Figure 14: Decoding Analysis using TimeCorr with variance of 1 time point



to show a marginal increase, a pattern not replicated by the results from the Forrest dataset.

We reason that the variation in decoding capabilities across different TimeCorr temporal resolutions may be attributed to different levels of focus revealing different information about the dynamics. As revealed in single subject TimeCorr testing, when the resolution is high (low Gaussian variance), the recovered correlation structures tend to have higher temporal accuracy, whereas low resolution TimeCorr (high Gaussian variance) can recover correlation changing dynamics with more stability. Therefore, when high resolution TimeCorr is used to decode the datasets, the ISFC-based brain information (high order brain connectivities) will contain more accurate information on the temporal correlation structure, thereby increasing time-point-to-time-point decoding accuracy. However, when the temporal resolution of TimeCorr is too high (Gaussian variance of 1 time point), TimeCorr will tunnel-vision on a very narrow neighborhood of time points for its functional connectivity calculations, and thus fail to incorporate enough information to recover any meaningful dynamic correlation patterns.

Lastly, when comparing between results under different stimulus conditions (intact, paragraph, word, resting) within the Pieman dataset, we can see that as the complexity and integrity of the stimulus decreases, the decoding accuracy decreases for all orders of functional connectivity. This makes sense as the human brain should intuitively produce more potent cognitive response to more coherent and meaningful stimuli (intact), weaker re-

sponse to scrambled and irrational stimuli (paragraph, word), and little to no response to no stimuli (resting). One anomaly exists in the results from the TimeCorr implementation with a Gaussian variance of 1 time point, where the decoding accuracy of the first order functional connectivity in the paragraph condition exceeded that of the intact condition. However, this behavior becomes understandable after taking into consideration the results from the TimeCorr implementation with a Gaussian variance of 10 time points: while the decoding accuracy for the first order functional connectivity of the paragraph condition only showed marginal decrease, the decoding accuracy of the first order functional connectivity of the intact condition decreased drastically when the TimeCorr Gaussian variance changed from 10 time points to 1 time point. As discussed previously, when the TimeCorr's temporal resolution is too high, it may fail to include enough global information to recovery meaningful dynmaic correlation structures. As the functional connectivity under the intact condition possesses more structure and coherency, it may be suffer greater deterioration than the functional connectivity from the paragraph condition when subjected to over-high temporal resolution, thus explaining the anomaly.

3.4 Multi-Level Mixture Analysis

To verify our hypothesis that the functional connectivity at each level represent non-overlapping sources of cognitively relevant information that could potentially complement each other, we designed the Multi-Level Mixture Analysis. This analysis constructs a mixture model incorporating information from multiple orders of functional connectivity, and contrasts its decoding capabilities with decoding accuracy of functional connectivities at each individual level.

Given a fMRI dataset for S subjects, each possessing T time points and V voxels, to construct the Multi-Level Mixture Model:

1. Choose the desired number of levels of functional connectivities L and the appropriate Gaussian variance for TimeCorr.

2. Use the TimeCorr Level Up method to obtain all level functional connectivites for each subject.

3. Randomly divide the subjects in to four equal groups A_1, A_2, B_1 and B_2.

4. For each group, use the Multi-Subject TimeCorr ISFC method to compute the inter-subject functional connectivity for each level

5. Compute correlation matrices for each level for A and B (i.e. for group A, correlate A1's features at each time point with A2's features at each time point; similarly for group B). This yields one correlation matrix for group A and another for group B, for each level (including level 0). So if $L = 10$, this process would yield 11 correlation matrices for A and another 11 for B.

6. Apply Fisher Z-transform (r2z) on all of the correlation matrices. The transformed correlation matrices are labelled $z_{A0}, z_{A1}, z_{A2}, ..., z_{B0}, z_{B1}, z_{B2}, ..., z_{B10}$, where z_{A0} represent the Fisher Z-transformed correlation matrix for level 0 of group A.

7. Use constrained optimization by linear approximation (COBYLA) on the group A correlation matrices to compute the optimal weight matrix w to be used for decoding, where w will be applied to the Fisher Z-transformed matrices in the manner of the following equation:

   $$decoding\_matrix\_A = z2r(w[0] * z_A0 + w[1] * z_A1 + w[2] * z_A2 + ... + w[10] * z_A10)$$

   where $z2r$ is the Inverse Fisher Z-transformation to convert the weighted sum into the its proper correlation form. Intuitively, this step seeks to find the w that maximizes the decoding accuracy using $decoding\_matrix\_A$.

8. Use the same z2r(weighted z-transformed sum) formula and the optimal weights from the previous step to compute decoding accuracy for group B

9. Output w and the group B's decoding accuracy

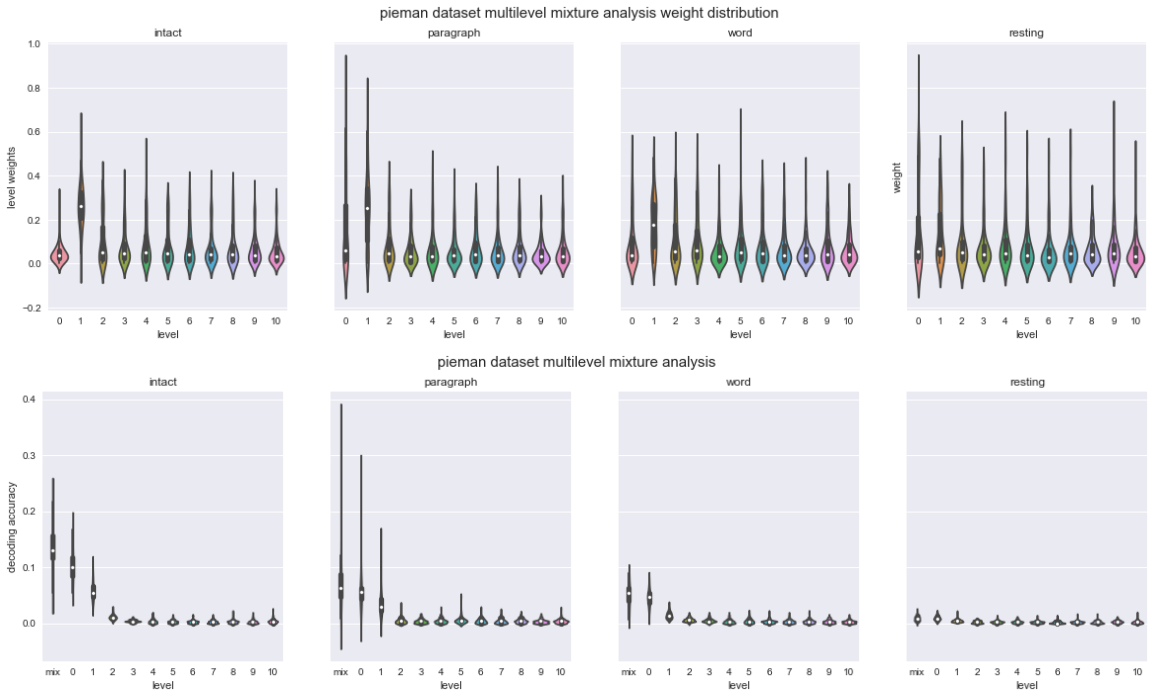
Three TimeCorr setups were implemented—each with Gaussian variances of 1000 time points, 10 time points, or 1 time point—to comprehensively understand the decoding capability of the mixture model and each individual level under varying resolutions. In addition, similar to the single level decoding analysis conducted previously, three different datasets were used for this analysis: Sherlock, Pieman, and Forrest. For each TimeCorr setup for

each dataset, we conducted 100 repetitions of decoding analysis with different random group divisions. All of the following results are the average of 100 repetitions.

Results:

1. The average results for each TimeCorr setup implemented on the Pieman dataset are displayed in Figure 15, 16, and 17.
2. The average results for each TimeCorr setup implemented on the Forrest dataset are displayed in Figure 18, Figure 19 and Figure 20.
3. The average results for each TimeCorr setup implemented on the Sherlock dataset are displayed in Figure 18, Figure 19 and Figure 21.

Figure 15: Multi-Level Mixture Analysis using TimeCorr with variance of 1000 time points



With each implementation of the Multi-Level Mixture Model, the subjects are divided into two equal groups, one for training and one for testing. As a result, we observe that training and testing using only half of the subjects caused the decoding accuracy to be a lot lower than when using the entire dataset. From a mathematical perspective, averaging

Figure 16: Multi-Level Mixture Analysis using TimeCorr with variance of 10 time points

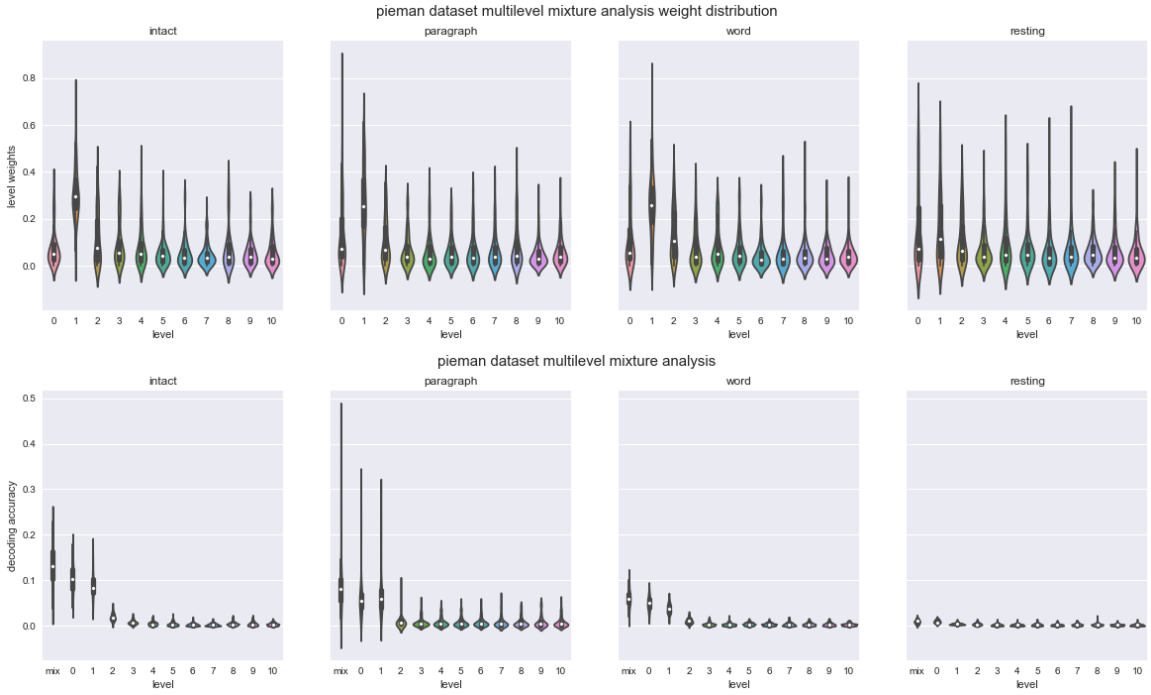


Figure 17: Multi-Level Mixture Analysis using TimeCorr with variance of 1 time point

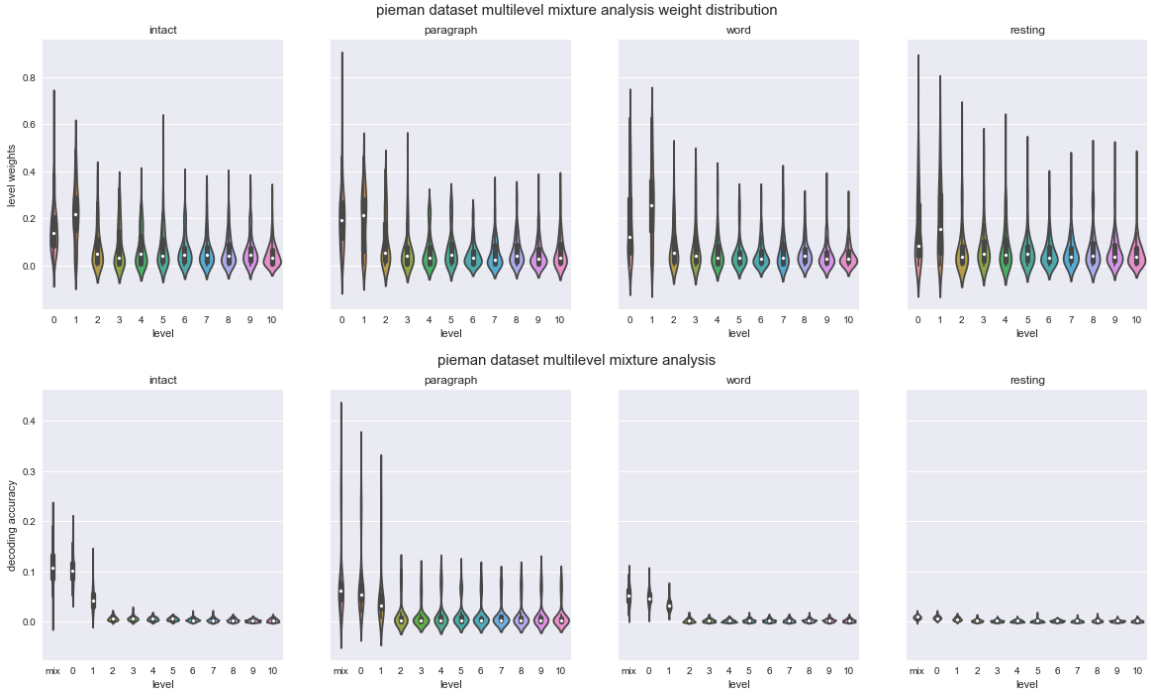


Figure 18: Multi-Level Mixture Analysis using TimeCorr with variance of 1000 time points

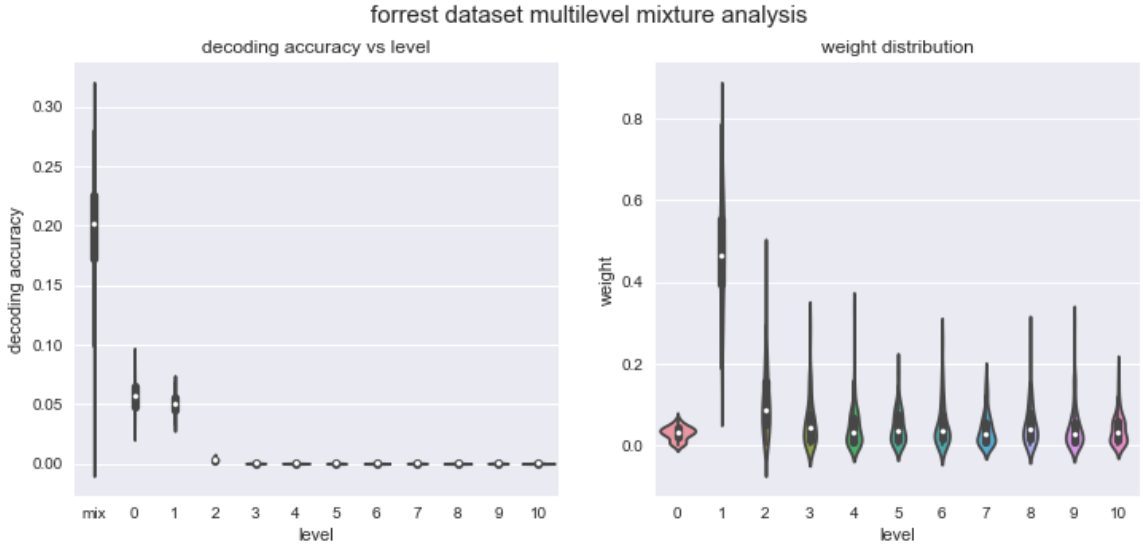
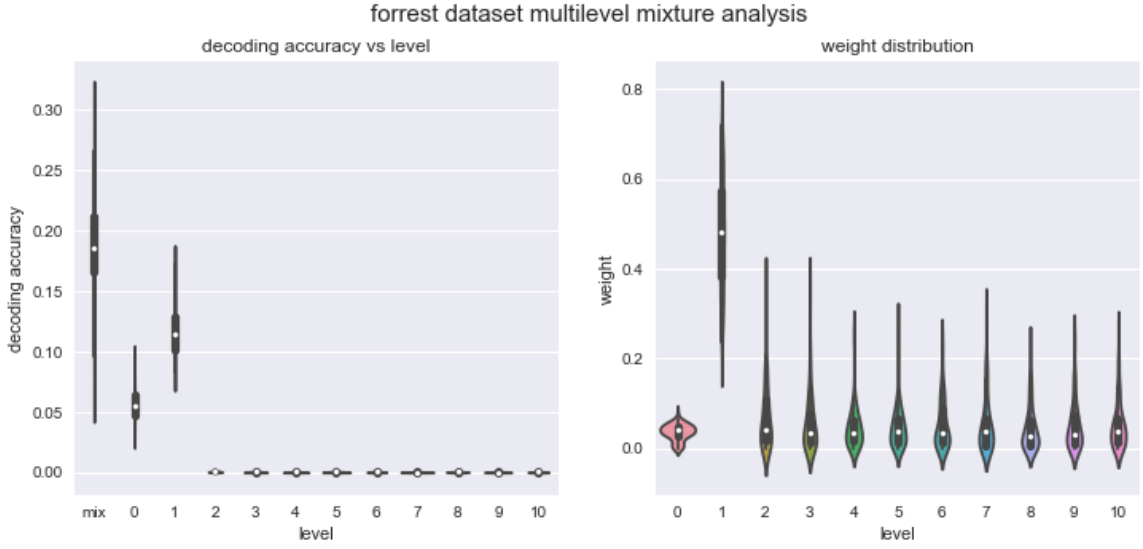


Figure 19: Multi-Level Mixture Analysis using TimeCorr with variance of 10 time points



over more subjects brings more stability to the dataset by smoothing out random noises in the dataset and adding emphasis to activities common to all subjects, which are typically stimulus-driven. Therefore, when less subjects are used in the calculation of inter-subject functional connectivity, the de-noising effect of averaging is weakened and less emphasis is put on activities that are stimulus-driven, thus causing lower decoding accuracy.

Figure 20: Multi-Level Mixture Analysis using TimeCorr with variance of 1 time point

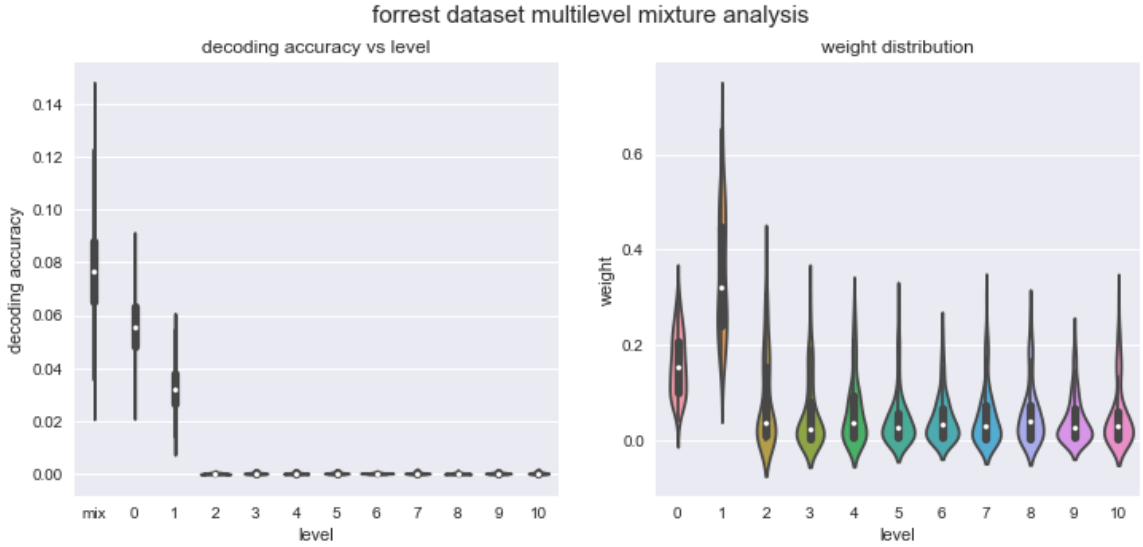
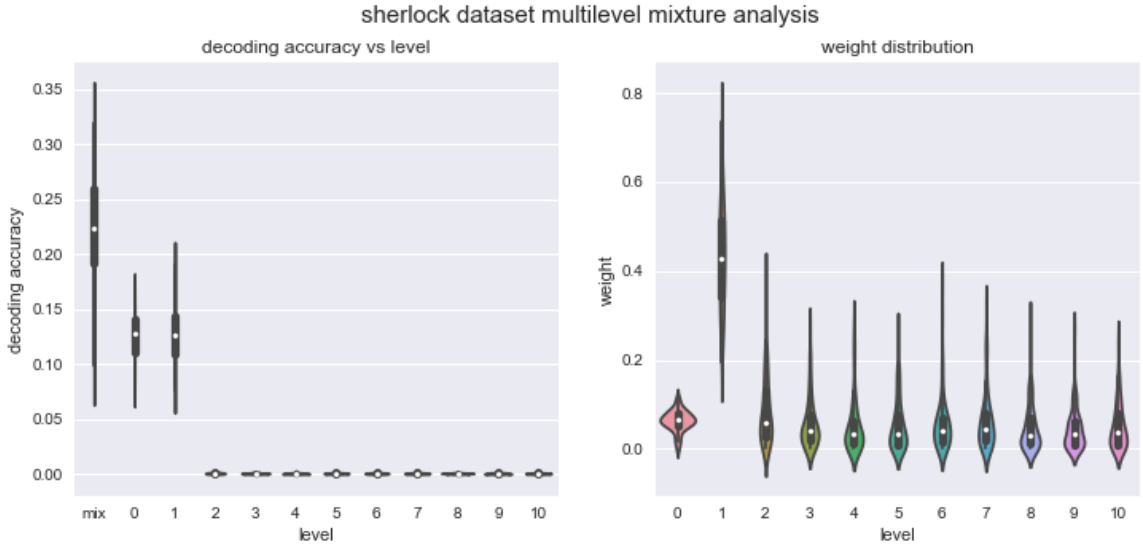


Figure 21: Multi-Level Mixture Analysis using TimeCorr with variance of 1 time point



One of the most important observations from the results of Multi-Level Mixture Analysis across all three datasets is that the Mixture Model produces significantly higher decoding accuracy than any single level of functional connectivity. When taking into consideration our hypothesis from the decoding analysis in the previous section, this is a very direct and compelling proof that different orders of functional connectivity contain non-overlapping information about brain dynamics and, although individual levels did not produce high

classification accuracy, when used in unison can greatly enrich data complexity and increase decoding capabilities.

The importance of each level is further confirmed by optimal weight distribution across all three datasets. While the first order functional connectivity always has the highest weight, the weight distribution between PCA reduced fMRI activation and higher order connectivities is almost uniform, indicating equal contribution to enrichment of stimulus-related information and improvement in decoding capability of the dataset. Additionally, we noticed that the optimal weight of the first order functional connectivity was significantly higher than the rest of the levels, suggesting that a greater amount of stimulus-driven brain dynamics information is contained in this level. The significance of the first order functional connectivity in the mixture model is also consistent with high resolution results in single level decoding analysis, where first level functional connectivity achieved the highest decoding accuracy among all levels.

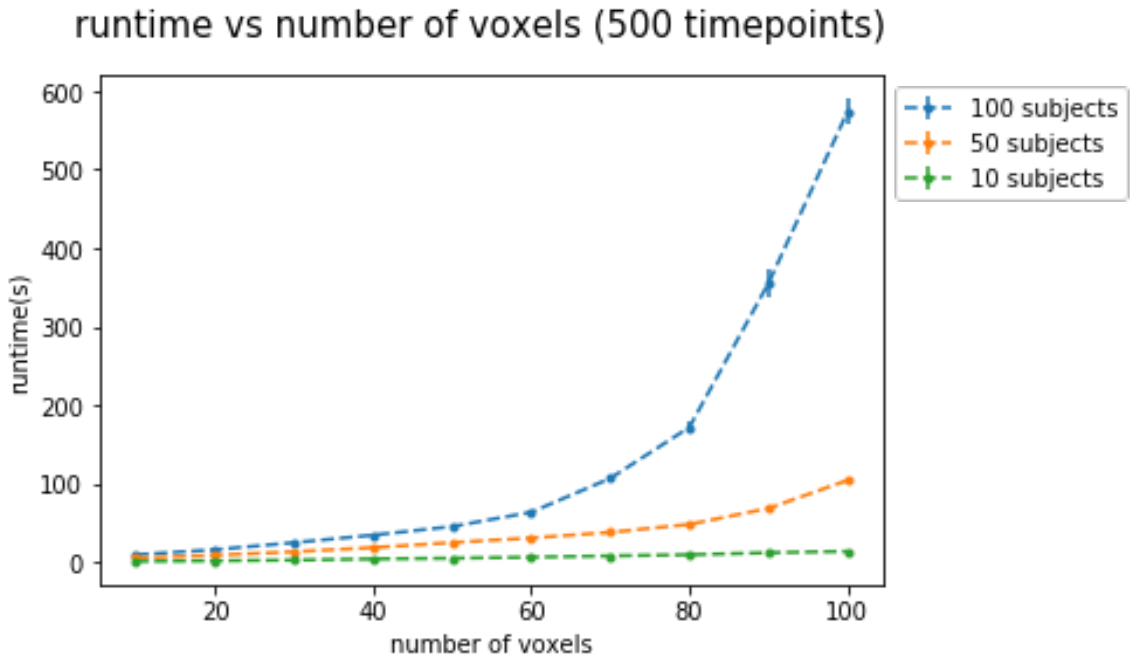
Like in the decoding analysis, we were curious if varying Gaussian variance (changing temporal resolution loss across levels) would influencing the results, therefore we repeated the Multi-Level Mixture Analysis on each dataset with multiple TimeCorr setups. However, the optimal weight distribution qualitatively remained the same in every implementation: the first level functional connectivity always possessed the highest weight. This further confirms that, although PCA reduced fMRI activations may sometimes produce higher decoding accuracy at low resolutions, first order functional connectivity actually contains the most information on brain dynamics that are stimulus-related.

3.5 High-Order Brain Dynamics Toolbox Benchmark Results

We wanted to make TimeCorr available to the public as a more effective alternative to the sliding window approach for functional connectivity related research, so we developed a Python toolbox that is available for public download through pip. Detailed documentation on how to use to toolbox is available online at (insert web address after completing documentation).

To make the TimeCorr-derived methods into a practical toolbox that can be widely used

by not only specialize laboratories but also the general public, we explored with different optimization and scaling methods to make it less resource intensive. At first, we explored with Cython, a optimizing static compiler that converts Python programs into C for faster implementation. However, we later discovered that multiprocessing has exponentially faster runtimes and scales well with large datasets. We also experimented with a combination of using Cython for preprocessing and multiprocessing for large scale calculation. However, on top of the low runtime achieved using multiprocessing, the time loss from C-Python transitions overshadowed the minimal speed gain from Cython. In the end, we decided to abandon Cython, and rely on multiprocessing as our main means of optimization.



Tests were performed on 2.6 GHz Intel Core i5 processor with 8GB 1600 MHz DDR3 Memory

The benchmark results of using TimeCorr to calculate inter-subject functional connectivity are shown in Figure 3.5. As we implemented multiprocessing as subject-wise parallel computing, we can see from the graph that runtime grows exponentially with the number of voxels, but linearly with the number of subjects. As starting every new process introduces time-loss, we decided that creating a process for each subject instead of for each voxel or time point would result in faster runtime.

4 Conclusion

1. application in medicine for classification data enrichment
2. Develop a lossless method to more accurately extract dynamic correlations from noisy data, breakthrough from the traditional sliding window approach
3. Due to the lossless nature of the method, we are able to calculate dynamic information at higher level
4. With richer information about the system, we are able to achieve higher decoding accuracy
5. Information from higher level helps with decoding, which can be applied in Criminal Investigation, etc....

Future work. PCA induced information loss. Explore Machine Learning alternatives to reducing data dimension. Autoencoder using Neural Network.

5 References

[1] Jeremy Manning, Xia Zhu, Theodore Willke, Rajesh Ranganath, Kimberly Stachenfeld, Uri Hasson, David M Blei, Kenneth A Norman. A probabilistic approach to discovering dynamic full-brain functional connectivity patterns. *bioRxiv* 106690, 2017

[2] Elena A. Allen, Eswar Damaraju, Sergey M. Plis, Erik B. Erhardt, Tom Eichele and Vince D. Calhoun. Tracking Whole-Brain Connectivity Dynamics in the Resting State. *Cerebral Cortex* March 2014;24:663–676

[3] Enrico Glerean, Juha Salmi, Juha M. Lahnakoski, Iiro P. Jaaskelainen, and Mikko Sams. Functional Magnetic Resonance Imaging Phase Synchronization as a Measure of Dynamic Functional Connectivity. *BRAIN CONNECTIVITY* Volume 2, Number 2, 2012

[4] Olaf Sporns, Giulio Tononi, Rolf Kotter. The Human Connectome: A Structural Description of the Human Brain. *PLOS Computational Biology* September 2005

[5] Erez Simony, Christopher J. Honey, Janice Chen, Olga Lositsky, Yaara Yeshurun, Ami Wiesel and Uri Hasson. Dynamic reconfiguration of the default mode network during narrative comprehension. *nature communications* July, 2016

[6] Elizabeth N. Davison, Benjamin O. Turner, Kimberly J. Schlesinger, Michael B. Miller, Scott T. Grafton, Danielle S. Bassett, Jean M. Carlson. Individual Differences in Dynamic Functional Brain Connectivity Across the Human Lifespan. *arXiv*:1606.09545v1

[7] Evelyn Tang and Danielle S. Bassett. Control of Dynamics in Brain Networks. *arXiv*:1701.01531v2

[8] Ankit N. Khambhati, Ann E. Sizemore, Richard F. Betzel, and Danielle S. Bassett. Modelling and Interpreting Network Dynamics. *bioRxiv* Apr. 4, 2017.

[9] Uri Hasson, Rafael Malach and David J. Heeger. Reliability of cortical activity during natural stimulation. *Cell Press* December 2009

[10] Roy Mukamel, Hagar Gelbard, Amos Arieli, Uri Hasson, Itzhak Fried and Rafael Malach. Coupling Between Neuronal Firing, Field Potentials, and fMRI in Human Auditory Cortex. *SCIENCE* VOL 309 5 AUGUST 2005

[11] Uri Hasson, Yuval Nir, Ifat Levy, Galit Fuhrmann and Rafael Malach. Intersubject Synchronization of Cortical Activity During Natural Vision. *SCIENCE* 12 MARCH 2004 VOL 303.

[12] Aya Ben-Yakova, Christopher J. Honey, Yulia Lerner and Uri Hasson. Loss of reliable temporal structure in event-related averaging of naturalistic stimuli. *NeuroImage* 63 (2012) 501–506

[13] Duha Al-Zubeidi, Mathula Thangarajh, PhDb, Sheel Pathak, Chunyu Cai, Bradley L. Schlaggar, Gregory A. Storch, Dorothy K. Grange and Michael E. Watson Jr.

[14] Gregory C. Burgess, Sridhar Kandala, Dan Nolan, Timothy O. Laumann, Jonathan D. Power, Babatunde Adeyemo, Michael P. Harms, Steven E. Petersen, and Deanna M. Barch. Evaluation of Denoising Strategies to Address Motion-Correlated Artifacts in Resting-State Functional Magnetic Resonance Imaging Data from the Human Connectome Project. *BRAIN CONNECTIVITY* Volume 6, Number 9, 2016

[15] Evan M. Gordon, Timothy O. Laumann, Babatunde Adeyemo, Adrian W. Gilmore, Steven M. Nelson, Nico U.F. Dosenbach, Steven E. Petersen. Individual-specific features of brain systems identified with resting state functional correlations. *NeuroImage* 146 (2017) 918–939

[16] Timothy O. Laumann, Abraham Z. Snyder, Anish Mitra, Evan M. Gordon, Caterina Gratton, Babatunde Adeyemo, Adrian W. Gilmore, Steven M. Nelson, Jeff J. Berg5, Deanna J. Greene, John E. McCarthy, Enzo Tagliazucchi, Helmut Laufs, Bradley L. Schlaggar, Nico U. F. Dosenbach, and Steven E. Petersen. On the Stability of BOLD fMRI Correlations

[17] Deanna J. Greene, Jessica A. Church, Nico U.F. Dosenbach, Ashley N. Nielsen, Babatunde Adeyemo, Binyam Nardos, Steven E. Petersen, Kevin J. Black and Bradley L. Schlaggar. Multivariate pattern classification of pediatric Tourette syndrome using functional connectivity MRI. *Developmental Science* 19:4 (2016), pp 581–598

[18] Christopher D. Smyser, Nico U.F. Dosenbach, TaraA. Smyser, AbrahamZ. Snyder, Cynthia E. Rogers, Terrie E. Inder, Bradley L. Schlaggar and Jeffrey J. Neil. Prediction of brain maturity in infants using machine-learning algorithms. *NeuroImage* 136 (2016) 1–9

[19] Maital Neta, XFrancis M. Miezin, Steven M. Nelson, Joseph W. Dubis, Nico U.F. Dosenbach, Bradley L. Schlaggar and Steven E. Petersen. Spatial and Temporal Characteristics of Error-Related Activity in the Human Brain. The Journal of Neuroscience, January 7, 2015

[20] Jonathan D. Power, Bradley L. Schlaggar, and Steven E. Petersen. Recent progress and outstanding issues in motion correction in resting state fMRI. *NeuroImage* 105 (2015) 536–551

[21] Nikos K. Logothetis, Jon Pauls, Mark Augath, Torsten Trinath and Axel Oeltermann. Neurophysiological investigation of the basis of the fMRI signal. *Nature* 412, 150-157 (12 July 2001)

[22] Joshua D. Greene, R. Brian Sommerville, Leigh E. Nystrom, John M. Darley and Jonathan D. Cohen. An fMRI Investigation of Emotional Engagement in Moral Judgment. *Science* 14 Sep 2001: Vol. 293, Issue 5537

[23] K.J. Friston, A. Holmes, C.J. Price, C. Buchela and K.J. Worsley. Multisubject fMRI Studies and Conjunction Analyses. *NeuroImage* Volume 10, Issue 4, October 1999

[24] K.J. Friston, P. Fletcher, O. Josephs, A. Holmes, M.D. Rugg and R. Turner. Event-Related fMRI: Characterizing Differential Responses. *NeuroImage* Volume 7, Issue 1, January 1998

[25] S. Ogawa, T. M. Lee, A. R. Kay, and D. W. Tank. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *PNAS* December 1, 1990

[26] Kenneth A. Norman, Sean M. Polyn, Greg J. Detre and James V. Haxby. Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Science* (2006)

[27] Nicholas B. Turk-Browne. Functional Interactions as Big Data in the Human Brain. *SCIENCE* (342) November 1st, 2013.

[28] M. Rubinov and O. Sporns. Complex network measures of brain connectivity: uses and interpretations. *NeuroImage*, 52:1059-1069, 2010.

[29] R. E. Betzel, J. D. Medaglia, L. Papadopoulos, G. Baum, R. Gur, R. Gur, D. Roalf, T. D. Satterthwaite, and D. S. Bassett. The modular organization of human anatomical brain networks: accounting for the cost of wiring. *Network Neuroscience*, page Advance online publication. doi:10.1162/netn.a.00002., 2017.

[30] R. C. Craddock, G. A. James, I. P E Holtzheimer, X. P. Hu, and H. S. Mayberg. A while brain fmri atlas generated via spatially constrained spectral clustering. *Human Brain Mapping*, 33(8):1914-1928, 2012.

[31] B. T. T. Yeo, F. M. Krienen, J. Sepulcre, M. R. Sabuncu, D. Lashkari, M. Hollinshead, J. L. Roffman, J. W. Smoller, L. Zollei, J. R. Polimieni, B. Fischl, H. Liu, and R. L. Buckner. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophsiology*, 106(3):1125-1165, 2011.

[32] Janice Chen, Yuan Chang Leong, Christopher J Honey, Chung H Yong, Kenneth A Norman, and Uri Hasson. Shared memories reveal shared structure in neural activity across individuals. *Nature Neuroscience* 20, 115–125, 2017

[33] Michael Hanke, Florian J. Baumgartner, Pierre Ibe, Falko R. Kaule, Stefan Pollmann, Oliver Speck, Wolf Zinke, and Jörg Stadler. A high-resolution 7-Tesla fMRI dataset from complex natural stimulation with an audio movie. *Scientific Data* Article number: 140003, 2014