MOTION CORRECTION FOR fMRI DATA USING CONDITIONAL TRANSITION REGIME SWITCHING GENERAL AUTOREGRESSIVE CONDITIONAL HETEROSKEDASTICITY MODELS

Teague R. Henry

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Psychology and Neuroscience.

Chapel Hill
2017

Approved by:

Kathleen Gates

David Thissen

Vladas Pipiras

Kelly Giovanello

Patrick Curran

ABSTRACT

TEAGUE R. HENRY: Motion Correction for fMRI data using Conditional Transition Regime Switching General Autoregressive Conditional Heteroskedasticity Models.
(Under the direction of Kathleen Gates)

This dissertation develops the Conditional Transition Regime Switching General Autoregressive Conditional Heteroskedasticity Model (CTRS-GARCH) for motion correction of functional magnetic resonance imaging (fMRI) region of interest (ROI) time series. This methodology brings together finite mixtures, hidden Markov modeling, proportional odds modeling, and multivariate volatility analysis to develop a non-destructive (in the sense it does not remove time points) method for removing the influence of motion from the correlation matrices that result from functional connectivity analyses on fMRI data. This dissertation develops the analytics and estimation procedures for the CTRS-GARCH, evaluates the performance using simulations, and uses the CTRS-GARCH to evaluate motion artifacts on an empirical dataset.

For my wife, Jennifer MacCormack, my colleagues, and Cornelius F. Todd. All played an important part in making this happen.

ACKNOWLEDGMENTS

I would like to acknowledge my advisor, Dr. Kathleen Gates, without whom this project would never have happened. Additionally, Zachary Fisher, Stephanie Lane, and Nathan Markiewitz who suffered through my endless explanations and questions. My committee, who managed to read a 70 page proposal without immediately refusing to sit on the committee. Finally, to my dear wife, Jennifer MacCormack, whose input and questions made this a significantly $(p < .05)$ stronger dissertation .

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 FMRI METHODOLOGY AND THE MOTION CONFOUND

Functional Magnetic Resonance Imaging, or fMRI, is a technical methodology that allows researchers to investigate so called *functional* hypotheses regarding brain activity. These hypotheses can be termed functional in nature as fMRI nominally measures an analog of neural activity in the brain. This differs from structural MRI, as structural MRI does not assess the active functioning of the brain, and rather provides an image of the anatomical structure. fMRI methodology has been used to study everything from the functioning of typical adults (Biswal, Yetkin, Haughton, & Hyde 1995) to task based performance of children with autism (Gomot et al. 2006), and a full review of the uses of fMRI is beyond the scope of this dissertation. Instead, this dissertation will focus on a particular issue of fMRI data collection that cuts across every application, that of motion of the participant when the fMRI data are being collected. Motion distorts the data, raising the chance of false positive findings, and is a particularly salient concern for applications regarding children and individuals with clinical disorders such as autism.

There are a variety of motion correction techniques, and this dissertation will review several of them, however the purpose of this dissertation is to introduce a new motion correction technique, the Conditional Transition Regime Switching Generalized Autoregressive Conditional Heteroskedasticity model (CTRS-GARCH), that is designed to target the particular issue of datasets with high motion where functional connectivity is the outcome under study. This method effectively breaks down the fMRI time series into a series of correlation matrices, and accounts for how motion affects the differences between these correlation matrices. For the substantive user of this motion correction method, the CTRS-GARCH can either provide a motion corrected correlation matrix for each subject, or a motion corrected time series depending on the needs of the researcher.

The outline of this dissertation is as follows. First, the method of fMRI will be reviewed to provide a grounding in the mechanics of data collection. Second, the problem of motion as well as existing motion correction methods will be reviewed. Third, time series methodology, regime switching and

generalized autoregressive conditional heteroskedasticity models (GARCH) will be reviewed. Fourth, the CTRS-GARCH model will be analytically derived and a method of estimation will be outlined. Fifth, a set of simulations will be presented comparing the CTRS-GARCH to other commonly used motion correction methods. Sixth, and finally, an empirical study will be presented, again comparing the performance of the CTRS-GARCH with other commonly used methods for motion correction.

1.1 fMRI Data Acquisition

fMRI stands for functional magnetic resonance imaging, and this technique of data acquisition seeks to acquire images that represent the functioning of various regions of the brain. While the physics of the data acquisition are complex, and a full treatment is beyond the scope of this dissertation, it is useful to provide a brief overview of how fMRI data are acquired, and in particular what data are acquired, as the actual signal required is not the true neurological functioning of the brain, but rather a proxy based on blood oxygen flow.

1.1.1 Magnetic Resonance Imaging

Magnetic resonance imaging, of which fMRI is a particular subtype, is a technology that allows researchers to form 3D images of solid objects, using an interaction between a strong magnetic field and radiofrequency pulses (Huettel, Song, & McCarthy 2014). The process by which these images are collected is fairly straightforward.

To begin, an MRI scanner contains a large superconducting electromagnet, used to generate a static magnetic field of anywhere between 1.5 Tesla to 11 Tesla, with the most common field strengths for fMRI on humans being 3T and 7T. This magnetic field aligns the hydrogen atoms in the human body, most of which are contained in the water that makes up the majority of the volume of the human body, into a low energy state (specifically, the static magnetic field aligns the spin states of protons in atoms to a parallel spin, which is lower energy). The MRI scanner also contains radiofrequency coils. Once the magnetic field has aligned the protons, the radiofrequency coils release a series of pulses of electromagnetic waves. These waves are absorbed by a percentage of the aligned protons (in the case of fMRI the aligned protons are concentrated in deoxygenated blood), which knocks these protons into a higher energy state, where the spin is not aligned to the magnetic field, and when the pulse ends, the protons that were in a high energy state fall (the analogy here of potential energy and an object

falling into a gravity well is appropriate) into the low energy aligned spin state. This "fall" releases an electromagnetic wave that can be detected by the radiofrequency coils (Huettel et al. 2014).

This release of energy has no spatial information because as the protons are uniformly aligned by the static magnetic field, and the radiofrequency pulse is also uniform across the subject. The final piece of MRI imaging is the gradient coils. These coils modify the magnetic field by emitting a series of pulses over the course of the scan (the *pulse sequence*) so that the magnetic field is no longer uniform across the entire subject. This lack of uniformity allows the scanner to calculate where a particular signal is coming from, so it is the gradient coils that allow MRI imaging to have the necessary spatial information.

The type of image, be it structural, functional, chemical or dynamic is determined by the pulse sequence of the gradient coils. By changing the pulse sequence, different parts of the subjects anatomy are aligned to the magnetic field differently. These differences in alignment, and the resulting differences in images are called *contrasts*. Utilizing different contrasts allows the researcher to obtain different types of images. For example, the T_1 contrast is used to examine the structural differences between grey matter and white matter in the brain (Huettel et al. 2014). The T_2 contrast picks up on areas that contain large amounts of fluids. For fMRI applications however, the most important contrast is the T_2^* contrast, which best shows the presence or absence of deoxygenated blood (Huettel et al. 2014).

Now, one might note that neuronal activity has not been mentioned as something directed measured by MR imaging. Indeed, actual neuronal activity, the neurons firing, is not measured by MRI in general, or fMRI in particular. Instead, fMRI presents the BOLD signal, standing for blood oxygenation level dependent signal. In the seminal paper for fMRI methodology, Ogawa, Lee, Kay, & Tank (1990), showed using rats that the T_2^* contrast showed the presence or absence of deoxygenated blood, and that a decrease in the amount of deoxygenated blood was correlated with increased neuronal activity in the area. Biologically, this makes sense as oxygenated blood carries both oxygen and glucose, and when neurons activate, they need that oxygen and glucose to continue to function. However, it is important to note that this BOLD response is only a correlate to neuronal activity in the area, and that neuronal activity and the BOLD response act on different time scales, with neurons

acting in milliseconds, while BOLD operates over seconds (Logothetis 2008). Moving forward, this distinction is important to keep in mind as the specifics of fMRI study design are discussed.

1.1.2 The BOLD time series

As the scanner is run on a subject the fMRI image is built up in a series of slices. These slices are produced as a product of the pulse sequence, and are combined, and then corrected, to form a voxel by voxel image of the subject's brain at a given time point. The *voxel* is the smallest unit of analysis in any MRI study. A *voxel* is a 3-dimensional space, that can be in the order of millimeters to a side, or centimeters, where per time point a single number representing the signal strength within that voxel is assigned. In the case of fMRI, this signal is the BOLD signal, and so the number assigned to each voxel at a time point represents the amount of deoxygenated blood in that voxel, and by proxy the neuronal activity in that voxel.

As was mentioned above, an MRI image is a 3-dimensional dataset. With an fMRI image, we have an additional dimension, that of time. In fMRI, BOLD response is collected per voxel per time point. The interval between time points can be anywhere between less than a second and more than five seconds, with a two second interval being a common choice. Smaller between time point intervals result in less spatial resolution, and a more noisy dataset, while longer time intervals increase spatial resolution, but of course decrease the temporal resolution. BOLD response after neuronal activity is of the order of several seconds (Huettel et al. 2014), making these time intervals consistent with the process under observation.

While the data that come directly out of a scanning session can be conceptualized as a multivariate time series, it is in fact unusable for analysis in its raw form. Raw data directly from the scanner are beset by gross motion artifacts, as well as a variety of physiological noise signals, which need to be removed before any analysis can be taken place. Before any analysis can be done the raw data must be preprocessed.

1.1.3 fMRI Preprocessing

Preprocessing MRI data is a complex endeavor, and like many aspects of MRI overall, a detailed treatment is beyond the scope of this dissertation. Aspects of preprocessing include quality assurance, where researchers examine the images to ensure that no scan artifacts are present, as well as the

application of various filters to remove physiological noise, such as the cyclical effect of breathing from the data (Huettel et al. 2014). The one aspect of preprocessing that will be discussed here is motion correction.

Motion is one of the most pernicious causes of data distortion in any MRI study (Power, Barnes, Snyder, Schlaggar, & Petersen 2011). A full treatment of the effects of motion on fMRI outcomes is presented later in this chapter, but for now we can focus on the more obvious effects of motion. The most obvious effect of motion is the shift in voxel location that occurs. Very small movements, such as 5mm, can result in the locations of the voxels being shifted dramatically (Huettel et al. 2014). This shift can cause significant differences in signal strength, and given that the head can and will move multiple times during a scanning session, the accumulated distortions can result in nearly unusable data.

To combat this, *coregistration* by *rigid-body transformation* is used (Huettel et al. 2014, Pg. 304). Coregistration refers to the aligning of two images, which is typically applied recursively to each time point, resulting in a final set of images that have been re-aligned to match one another. Rigid-body transformation is the process of performing the realignment, which is made simpler due to the fact that the head and brain do not change shape over the course of the scan. In order to perform this realignment, the motion parameters, namely movement in all three directions, as well as rotation in all three dimensions need to be estimated. There are a variety of techniques with which to estimate these parameters, and a different technique is used for each popular preprocessing software package. For example, FSL (FMRIB (Functional MRI of the Brain) Software Library) uses an algorithm termed MC-FLIRT (Motion Correction - FMRIB's Linear Image Registration Tool); (Jenkinson, Bannister, Brady, & Smith 2002), while SPM (Statistical Parametric Mapping) uses an unnamed coregistration technique (Penny, Friston, Ashburner, & Nichols 2006).

While realignment, as we will see, does not remove all motion artifacts, it does reduce the gross issues of changing voxel positions. Once coregistration, filtering and time slice correction is performed, the multivariate time series of the BOLD response on a voxel by voxel level is finally ready for analysis.

ROI selection

As an important aside, it is often the case that a researcher does not want to work with voxel level data. Regions of Interest (ROI) analysis for fMRI data in particular can provide an alternative. ROIs are sets of (usually) spatially contiguous voxels that are aggregated in some way. This aggregation can be done using a simple mean, or a more complex spatial weighting scheme, but the idea remains the same, and that is that the aggregate signal from the set of voxels is a more reliable measure of functional connectivity in that region. The sets of voxels that make up each ROI can be determined from a preexisting atlas, such as the Automated Anatomical Labeling atlas (AAL; Tzourio-Mazoyer et al. (2002)), the Harvard-Oxford atlas (HO; Desikan et al. (2006)) or the canonical Brodmann atlas Brodmann (1909). Alternatively, ROI maps can be generated on a per subject basis by examining functional connectivity between voxels and grouping voxels accordingly, or by using a cognitive task and determining the set of voxels that activate during that task. There are several algorithms for doing this, one example is the Craddock atlas generator (Craddock, James, Holtzheimer, Hu, & Mayberg 2012).

1.1.4 fMRI Outcome Types

While the true output of an fMRI session is a multivariate time series of the estimated BOLD response, we can divide the outcomes of interest into two broad camps, that of *site based activation*, or that of *functional connectivity*. Site based activation refers to the presence or absence of a signal at a particular region of interest, and is used to infer activation of that brain region. For example, if a participant is given a visual stimuli task, the V1 area of the primary visual cortex shows a increased BOLD response in that region. This would suggest that there is neurological activity in that region. While these outcomes have been of interest in the past, and are indeed still of interest for many task based studies, in this dissertation we focus on functional connectivity.

Functional connectivity attempts to measure coactivation of regions within the brain. This can be done with seed correlation (Huettel et al. 2014), where an initial seed voxel is designated, and all other voxels are then correlated with that seed voxel, to determine the functional connectivity between the seed and other areas. Alternatively, if one is working at a larger spatial scale, one can calculate a correlation matrix between all regions of interest (ROIs), which can establish a functional

connectivity map.

The covariance matrix, or the equivalent correlation matrix of the pairwise relations between ROIs is the most common data analysis object in functional connectivity analysis. While there are several methods that use the raw time series to attempt to construct a causal relation network, such as SEM approaches (J. Kim, Zhu, Chang, Bentler, & Ernst 2007) or dynamic causal modeling (Friston, Harrison, & Penny 2003), the most common approach to functional connectivity is to use the observed time series to calculate the pairwise correlation matrix, possibly use this matrix to calculate the partial correlation matrix, and to treat these matrices as the "network" of functional connectivity (Smith, Menon, & Thompson 2011). From this network, a variety of network statistics can be calculated and inferential procedures can be performed (Bullmore & Sporns 2009). This reliance on the correlation matrix for functional connectivity requires the ability to robustly estimate the true correlation matrix of the fMRI signal, uncontaminated by any sort of noise process.

Of particular note is the idea of dynamic functional connectivity (Hutchison et al. 2013). In previous years, functional connectivity has been calculated in a static fashion, treating whatever measure of functional connectivity calculated as the same across an entire block of time within the overall acquisition. Recently there has been a movement towards a more dynamic view of functional connectivity as changing over the course of acquisition. Methods to assess this include sliding windows (Hindriks et al. 2016), multivariate GARCH models (Lindquist, Xu, Nebel, & Caffo 2014), independent vector analysis techniques (Ma, Calhoun, Phlypo, & Adali 2014) and state-space modeling (Molenaar, Beltz, Gates, & Wilson 2016).

1.1.5 fMRI Study Design

Broadly, fMRI studies can be divided into two types, *task-based* and *resting state*. Task based study designs were the original design for fMRI studies and are, for example, used in studies regarding memory and cognition. A typical task based fMRI design consists of dividing the acquisition period into a set of activities, possibly repeating activities across acquisitions, or having a unique set as the study demands. The precise timings of these tasks are noted, and used to divide the acquired fMRI signal into blocks if the tasks are not interleaved, or to parse out which time points should be assigned to each task if the tasks are. These task based designs are typically hypothesis driven

(Huettel et al. 2014), where researchers have an *a priori* idea of the regions or voxels that are active during a particular task, and can describe a hypothesis that then can be tested.

Resting state data, which originated in the seminal paper of Biswal et al. (1995), involves having the subject simply rest in the scanner. Resting state data can provide a look at the general functional connectivity of a subject, and is often the data used to examine between-subject differences in functional connectivity. That being said, resting state fMRI is a data-driven affair, where there is a typical lack of hypotheses regarding the data (Huettel et al. 2014).

1.2 Motion as a Confound

One of the central themes of this dissertation is that motion has wide ranging and disruptive effects on fMRI data. As was mentioned above, motion during a scan physically changes the positioning of the voxels. During preprocessing this is corrected using coregistration and realignment. However, motion also disrupts the homogeneity of the magnetic field in the MRI scanner, and the effects of motion cannot be completely corrected using realignment.

This issue was first presented in the neuroimaging literature in 2011 and 2012 by three independent research groups, Power, Barnes, et al. (2011), Satterthwaite et al. (2012) and van Dijk, Sabuncu, & Buckner (2011). The general conclusion from all three of these groups is that small amounts of motion cause systematic differences in functional connectivity, that could be misinterpreted as between group differences, or other substantive findings. Specifically, Power, Barnes, et al. (2011) found that spatially distant regions had their correlation reduced by movement, while spatially close regions had their correlation increased by movement, and this general effect had a large degree of variance both within and between subjects. Both Satterthwaite et al. (2012) and van Dijk et al. (2011) had similar findings to Power, Barnes, et al. (2011). Furthermore, this distortion in the fMRI signal was found to occur during and immediately after the movement. As Power, Barnes, et al. (2011) illustrates with an example, an individual with a single severe movement will have better quality data overall than an individual with continuous micromovements, or small movements that occur in a matter of seconds, with individuals occasionally returning to the place of origin. It is hard to overstate the severity of the motion artifact. Power et al. (2014) states that the artifact caused by even small amounts of motion could result in spurious findings, and several studies have already been reevaluated to ensure that

their findings are robust to the motion artifact (e.g. Fair et al. (2012)). Figure 1.1 below shows the time series of 5 ROIs from an adolescent with autism spectrum disorder, along with the time series of frame displacement, a measure of motion that will be explained later in this dissertation.

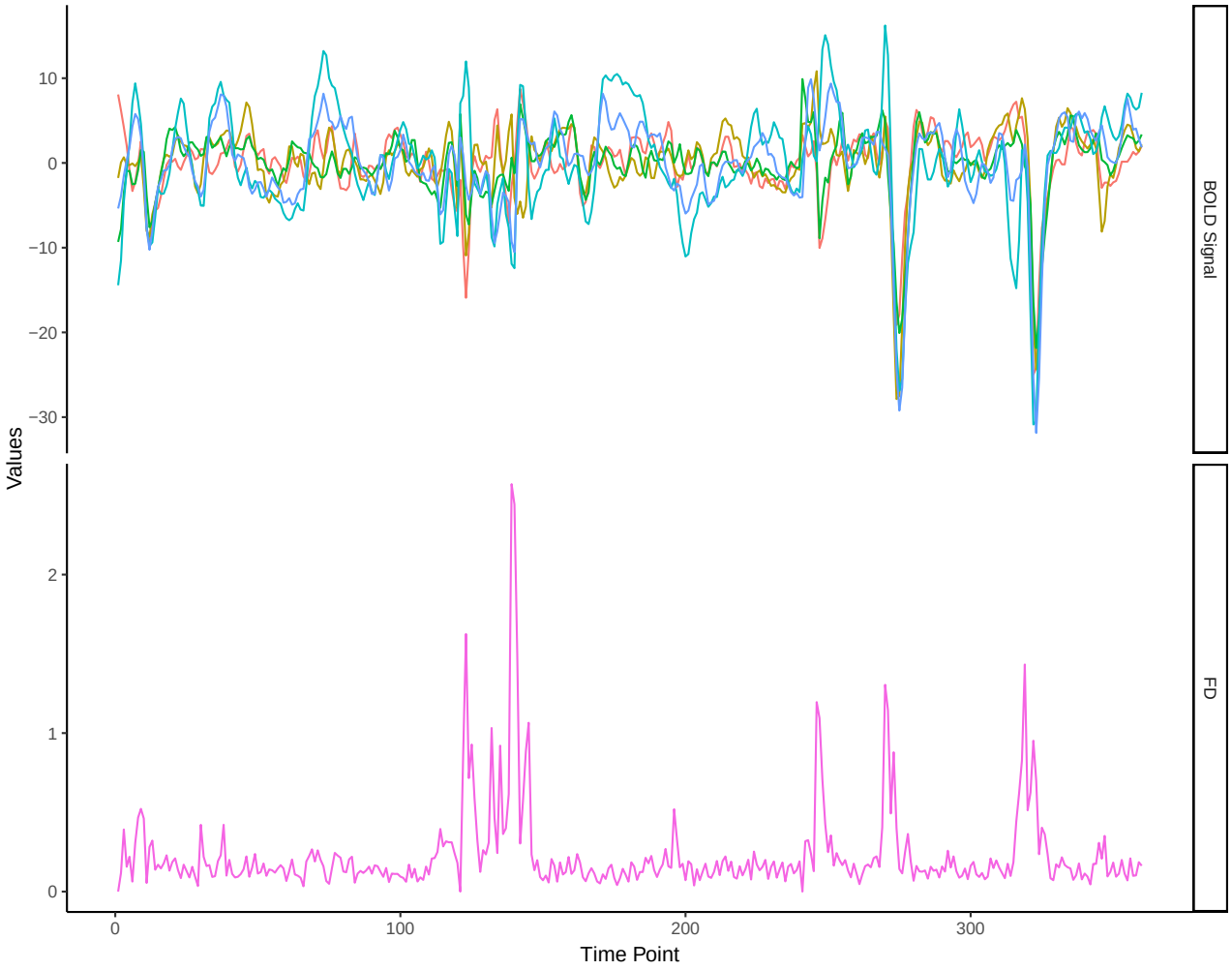


Figure 1.1: Timeseries of BOLD Signal and Frame Displacement for an adolescent with autism. higher values of frame displacement indicates higher amounts of motion. Note the large drop in BOLD signal for two of the occasions of motion, as well as the increase in volatility during the large amount of motion during the middle of the scan.

From the figure above, there are three important aspects of the motion artifact to note. The first is that the motion artifact does change the correlation structure. Note for the large decreases in BOLD, those time series are likely correlated at near 1. Note too the change in variance. Finally, this figure shows the heterogeneous nature of the motion artifact, in that motion does not have same effect on the BOLD signal each time.

With the problems presented by motion artifacts so revealed there has been an increased interest in motion correction techniques that can be used in tandem with the rigid-body correction already present in preprocessing. In the following section, we examine in detail the two most popular methods, 24-parameter regression and scrubbing.

1.2.1 Motion Correction Techniques

24 Parameter Motion Regression

24 parameter motion regression was originally proposed as a solution to the motion artifact by Friston, Williams, Howard, Frackowiak, & Turner (1996). This regression takes the 6 motion parameters, translation in the X, Y and Z direction, as well as the pitch P, roll R and yaw W for rotation, and uses them and time lagged expansions to attempt to regress the effects of motion out of the fMRI signal. The 24 parameter regression which incorporates 2 time points (Power et al. 2014) is as follows:

$$S_t^* = S_t - \hat{S}_t \tag{1.1}$$

$$\hat{S}_t = \beta_0 + \beta_1 X_t + \beta_2 X_t^2 + \beta_3 X_{t-1} + \beta_4 X_{t-1}^2 + \beta_5 Y_t + \cdots + \beta_2 4 W_{t-1}^2. \tag{1.2}$$

Where S_t^* is the corrected signal for the S voxel and S_t is the original value for the S voxel at time point t. While at first blush, this technique appears to be promising for extracting the motion artifact, evaluations of its performance suggest that it is highly inefficient at removing the effect of motion (Power, Barnes, et al. 2011; Power et al. 2014). Additionally, increasing the time lagged information from one lagged time point to two and three does improve the extraction of the motion artifact, but only slightly (Power et al. 2014). This lack of performance is likely due to the assumption of homogeneity inherent in the regression approach. In this approach, motion is considered to have the same effect on each voxel. However, this is likely not the case, as the effect of motion on alignment potential is not homogeneous (Power, Barnes, et al. 2011; Huettel et al. 2014).

It is also worthwhile to note that a common regression technique used in preprocessing, global signal regression, is highly effective at reducing the effect of motion (Power et al. 2014). However, global signal regression has the potential to distort the fMRI signal independently of motion. Murphy,

Birn, Handwerker, Jones, & Bandettini (2009) and Saad et al. (2012) indicate that the use of global signal regression can induce negative correlations between voxels. The efficacy of global signal regression is an ongoing area of research, however, given that it is not a dedicated motion correction technique that potentially has additional consequences suggests that it is not an ideal option for correcting the effect of motion.

Theoretically, regressing the effect of motion out of the time series is an attractive option from several views. The first being that the regression preserves the temporal nature of the time series. As opposed to scrubbing, the next correction technique to be discussed, regression does not remove time points wholesale, but rather attempts to remove the portion of signal that motion account for. If this was successful, it would leave an intact time series, so that both static functional connectivity analysis could be conducted, but more importantly dynamic functional connectivity (which ideally needs an intact time series) could be performed. Furthermore, the use of regression on high motion populations, such as pediatric populations, or individuals with clinical disorders such as autism or ADHD would be ideal, as it would not result in the wholesale loss of time points that scrubbing results in. However, as regression is not effective at removing the motion artifact, this approach is not valid, but the advantages of a regression approach should be kept in mind as we continue the discussion of motion correction techniques.

Scrubbing

Scrubbing, or the wholesale removal of points above a particular threshold of motion, has been shown to be effective at significantly reducing the motion artifact (Power, Barnes, et al. 2011; Power et al. 2014; Satterthwaite et al. 2013a). Typically, to use scrubbing a single value that represents the aggregate level of movement at a given time point is calculated. While a mean of the motion parameters can certainly be used, a popular alternative is that of frame displacement. Frame displacement has several definitions, but one of the most widely used comes from Power, Barnes, et al. (2011), as follows:

$$FD_t = |\Delta X_{t:t-1}| + |\Delta Y_{t:t-1}| + |\Delta Z_{t:t-1}| + |\Delta P_{t:t-1}| + |\Delta R_{t:t-1}| + |\Delta W_{t:t-1}|. \quad (1.3)$$

Where the motion parameters are as before, and $\Delta_{t:t-1}$ is the difference between the values of a given parameter from $t-1$ to t. Frame displacement provides an aggregate look at the magnitude of the movement from one time point to the next. By convention, FD_1 is 0. As a measure of motion, frame displacement is appropriate given the lack of efficacy of the regression approaches to motion correction.

Once frame displacement has been calculated, scrubbing is a simple matter of removing time points based on some rule involving frame displacement. A common choice is a simple threshold, where any time points that exhibit over, for example, a frame displacement of .2mm for a fairly stringent threshold, or a frame displacement of .5mm for a more liberal threshold (Power, Barnes, et al. 2011). Another approach is to remove both the time points that exceed the threshold, and a number of time points afterwards. The reasoning of this is that the effect of motion is not only apparent at the time point of the movement, but also in the time points that follow, often for several tens of seconds (Power et al. 2014). An alternative to frame displacement is that of DVARS, which measures the change in the variance of an fMRI signal (Smyser et al. 2010). It is important to note that DVARS is not a measure of motion, but rather is a measure of the stability of the fMRI signal. As motion does greatly effect the variance of the fMRI signal, DVARS is related to motion, but could also be confounded with a dynamic fMRI signal that changes systematically.

Several studies have shown the efficacy of the scrubbing approach. Power et al. (2014) shows that when applied to cohorts of adults who exhibited different levels of mean FD, and had significant differences in their functional connectivity (which, theoretically, they had no reason to have), scrubbing could reduce those differences down to below significance. Satterthwaite et al. (2013b) uses spike regression, which is functionally equivalent to scrubbing, and can be implemented simultaneously to a regression approach. Their approach showed a marked improvement in the motion artifact.

While scrubbing is undoubtedly efficient at removing the motion artifact, there are a couple inherent problems with the approach. First, and most obviously, scrubbing removes time points, which forces researchers to use the reduced dataset, as imputation is not an option given that there is little information as to a correct imputation model. The use of the reduced dataset leads to reduced power for any statistical test, just on the basis of sample size reduction. Second, in populations with high

movement such as children and individuals with clinical disorders such as ADHD or autism, the number of time points removed could render a subject unusable, and in the case of many clinical disorders each subject is a considerable investment on the part of the researcher.

On the other hand, scrubbing has several advantages over other approaches. The first is that it is inherently non-parametric. The functional form of the motion artifact is not a factor in scrubbing, as any time points that exhibit excess motion are removed. The second advantage is that it provides a very clear picture of which time points are problematic.

1.2.2 Summary

Motion is a particularly salient issue in the use of fMRI data. Motion artifacts can cause significant changes in the functional connectivity structure, and can result in significant and spurious between group differences. To combat this, a variety of motion correction techniques have been proposed. The most popular of these techniques are nuisance regression and scrubbing. While 24 parameter motion regression has several desirable properties theoretically, its effectiveness at removing the motion artifact remains disputed (Power, Barnes, et al. 2011). Scrubbing on the other hand has several undesirable properties, but is effective at removing motion artifacts. This lack of a motion correction technique that can be used on a high motion population that will not remove time points wholesale motivates the methodological development in this dissertation. In the following chapter, the CTRS-GARCH model is developed. This motion correction technique non-parametrically and non-destructively detects and removes motion artifacts at the ROI level, making it an ideal method for use on high motion populations, where conserving data is a top priority.

2 THE CTRS-GARCH MOTION CORRECTION METHOD

2.1 Time Series and fMRI

While functional connectivity analyses typically result in a correlation matrix as output, fMRI data at its core is a multivariate time series. A time series, simply put is any set of data that is indexed by time, and has some sort of ordered inter-dependency. While this dissertation deals with a type of model that acts on time dependent variances, it is useful as a grounding to briefly describe the two most common types of dependence in univariate time series modeling, and to discuss difficulties with the multivariate generalization of these dependences. After this overview, we will move to a discussion of the generalized autoregressive conditional heteroskedasticity model (GARCH; Engle 1982), which forms the core of the model developed in this dissertation.

2.1.1 Autoregressive and Moving Average Time Series

The most common forms of dependency modeled in time series data are those of *autoregressive* effects and *moving average* effects (Hamilton 1994). Consider the following: y_t is indexed by t, with $t \in 1, \ldots, T$. An autoregressive effect of lag 1 (with a 0 intercept) would be described by:

$$y_t = \beta y_{t-1} + \epsilon_t. \tag{2.1}$$

Where β is the autoregressive effect, and ϵ_t are i.i.d error terms with some distribution (typically normal).

A moving average model (again, with a 0 intercept) would be described as :

$$y_t = \theta \epsilon_{t-1} + \epsilon_t. \tag{2.2}$$

Where θ is the moving average effect.

Both of these model specifications have multivariate extensions that for the most part will not be

discussed here. Of importance however is the rapid increase in the number of parameters to estimate in a wholly unrestricted multivariate AR or MA type model. For example, in an autoregressive lag 1 model that is bivariate, there are four potential lagged effects, the autoregressive effect of each variable on its next time point, and the cross lag effect of each variable on each other. This model complexity rapidly increases as more dimensions or lags are added to the data.

With the basic form of dependence so described, we can move onto a discussion of GARCH models, which form the core of the CTRS-GARCH model developed in this dissertation.

2.1.2 Univariate GARCH

The generalized autoregressive conditional heteroskedasticity model or GARCH, was first introduced by Bollerslev (1986) as a model to investigate returns on stock prices. Fundamentally, instead of examining the change in the expected value of a process, the GARCH model examines changes in the variance of a process, and is related to a moving average process. GARCH type models are useful when the pattern of change one expects to see in a process results in heteroskedastic errors, or changes in the variance of the process over time. Expected value models often make the assumption of conditional homoskedasticity, and so would be inappropriate for processes with changing variances or covariances. The specification for a GARCH(1,1) model is as follows:

$$y_t = \sqrt{h_t}\epsilon_t \tag{2.3}$$

$$h_t = \alpha + \beta h_{t-1}\epsilon_{t-1}^2 + \theta h_{t-1}. \tag{2.4}$$

Where h_t is the variance of the time series at time t, β is the effect of the empirical residual at lag, and θ is the effect of the estimated lagged variance ϵ_t has a variance of 1. Engle (2001) describes the GARCH model as modeling variance from a weighted average of the past squared residual values, where the weights decline with time delay but never fully go to 0. This forces a dependence structure where no matter the lag, the variances of two time points are dependent.

2.1.3 Multivariate GARCH

Typically, the goal in fMRI for functional connectivity is to perform analysis on the correlation or covariance matrix of the fMRI time series, given the assumption that these matrices map onto the functional relations between regions of interest. As a covariance is a generalization of the variance term for multivariate distributions, a multivariate GARCH is ideal for the analysis of fMRI data. The general specification is similar to a univariate GARCH model:

$$\mathbf{y}_t = \mathbf{H}_t^{1/2}\boldsymbol{\epsilon}_t \tag{2.5}$$

Where $\mathbf{y}_t$ is a p dimensional random variable, $\boldsymbol{\epsilon}$ is a p dimensional random variable with a covariance matrix of $\mathbf{I}_p$, and $\mathbf{H}_t$ is a $p \times p$ positive definite matrix, that acts as the covariance matrix of $\mathbf{y}_t$.

One of the most general specifications of a multivariate GARCH model is the vector error correction or VEC model of Bollerslev, Engle, & Wooldridge (1988), in which each element of the $\mathbf{H}_t$ matrix is a linear function of the lagged squared error terms and cross products, as well as the lagged elements of $\mathbf{H}_t$. The specification is as follows:

$$\mathbf{H}_t = \boldsymbol{\omega} + \boldsymbol{\alpha}\text{vech}(\mathbf{H}_{t-1}) \odot \text{vech}(\boldsymbol{\epsilon}_{t-1}\boldsymbol{\epsilon}_{t-1}') + \boldsymbol{\beta}\text{vech}(\mathbf{H}_{t-1}). \tag{2.6}$$

Where $\text{vech}(\cdot)$ is a vector stacking operation that operates on the upper triangular portion of a matrix, $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are square matrices of parameters of order $p(p-1)/2$, and $\boldsymbol{\omega}$ is a $p(p-1)/2$ vector of parameters. This model is not computationally tractable for dimensions of over 2 or 3, so restrictions must be put in place. One model with such restrictions that deserves a description in this dissertation is the dynamic correlation model (DCC) of Engle (2002). This model has been used in fMRI analysis by Lindquist et al. (2014), and indeed is the first and only GARCH type model to have found use on fMRI data.

The DCC is a highly constrained multivariate GARCH that decomposes the covariance matrix of

the time series as follows:

$$\text{Cov}(\mathbf{y}_t) = \mathbf{H}_t. \tag{2.7}$$

$$\mathbf{H}_t = \mathbf{D}_t \mathbf{\Gamma}_t \mathbf{D}_t. \tag{2.8}$$

Where $\mathbf{D}_t$ is an $p \times p$ diagonal matrix with the standard deviations of each dimension of the time series on the diagonal, and $\mathbf{\Gamma}_t$ is a correlation matrix.

The DCC then proposes that each diagonal element of $\mathbf{D}_t^2$, the matrix that contains the variances of each dimension, follows a univariate GARCH model with the following specifications:

$$\mathbf{D}_t^2 = \text{diag}(\boldsymbol{\omega}) + \text{diag}(\boldsymbol{\alpha}) \odot \text{diag}(\mathbf{y}_{t-1}\mathbf{y}_{t-1}') + \text{diag}(\boldsymbol{\beta}) \odot \mathbf{D}_{t-1}^2. \tag{2.9}$$

Where now $\boldsymbol{\omega}, \boldsymbol{\alpha}, \boldsymbol{\beta}$ are all $p \times 1$ vectors of parameters, diag is a function that when applied to a vector turns that vector into a diagonal square matrix with 0s off-diagonal, and when applied to a square matrix, returns only the diagonal elements in a diagonal matrix. $\odot$ is the Hadamard product, or elementwise multiplication. At the level of individual variances, the expression is as follows:

$$d_{(p)t}^2 = \omega_{(p)} + \alpha_{(p)} d_{(p)t-1}^2 \epsilon_{(p)t-1}^2 + \beta_{(p)} d_{(p)t-1}^2 \tag{2.10}$$

DCC is a univariate GARCH type model for each individual correlation in $\mathbf{\Gamma}_t$. Each off-diagonal element in $\mathbf{\Gamma}_t$ depends on the estimated value of that element $\mathbf{\Gamma}_{t-1}$ and the observed cross residual element in $\boldsymbol{\epsilon}_{t-1}\boldsymbol{\epsilon}_{t-1}'$. This reduces the total number of parameters to be estimate to $2p^2 + p$. The dual autoregressive processes for both the variances and the correlations lead to a smoothly varying covariance matrix over the time course. However, the lack of cross lagged effects is less than ideal for the context of fMRI. It is likely that changes in the functional connectivity of 2 regions could have effects on the functional connectivity of 2 different regions, and the DCC would fail to capture that for larger fMRI datasets. The one application of the DCC to fMRI, done by Lindquist et al. (2014), tested the approach on a bivariate relationship, in which case the DCC model fully captures the autoregressive effects on both the variances and the single correlation, and does not take into

account cross lagged relations between correlations. As this dissertation focuses on the effect that motion has on the whole functional connectivity network, an effect that is thought to be complex and not easily parameterized, a different approach than the DCC would be needed. This approach can be found in the regime switching multivariate GARCH model of Pelletier (2006).

2.1.4 Regime Switching Multivariate GARCH

The two previously discussed models, the VEC of Bollerslev et al. (1988) and the DCC of Engle (2002) both proposed smoothly varying correlations and variances across time. However the cost of those proposals was untenable complexity in the case of the VEC model, or harsh restrictions in the case of the DCC. In the application of this dissertation, the concern is not to completely reproduce the complex data generating process that results from motion artifacts, but to rather approximate the effect of motion on the underlying *true* functional connectivity correlation matrix to some high degree. This approximation can be achieved with a regime switching multivariate GARCH. This approach is, in brief, a Markov processed finite mixture modeling approach that allows the effect of motion to occur not as a continuous process, but rather as a series of shifts in the correlation matrix. These shifts can then persist over several time points. This reflects the persistent effect of motion on functional connectivity after motion has ceased, as well as reflecting the notion that an individual likely is in a different mental state while moving. Before we go into the details of the regime switching multivariate GARCH, the idea of regime switching will be briefly explained.

2.1.5 Regime Switching

Regime switching in time series analysis is a general concept that can be applied to almost any time ordered data set and model (C.-J. Kim & Nelson 1999). Regime switching can be illustrated with a mean switching autoregressive model:

$$y_t = \mu_t + \beta y_{t-1} + \epsilon_t. \tag{2.11}$$

$$\mu_t = \mu_{S_t=s}. \tag{2.12}$$

$$\epsilon_t \sim N(0, \sigma^2). \tag{2.13}$$

Where S_t is a discrete regime state, with possible values $s = \{1, \ldots, K\}$, and μ is the mean of

the process that changes with the state. In the vast majority of applications of regime switching the state at any time point is not known. In the model described in Equations 2.11-2.13, the mean of the univariate time series exhibits a regime switching structure, which is to say that during a particular state, the mean of the process does not change, but when the state transitions to a different value the mean then changes. In general, transition from state to state, or the regime switching is a first order Markov process (C.-J. Kim & Nelson 1999), which means:

$$P(S_t = s|y_1, \ldots, y_t, S_1, \ldots, S_{t-1}) = P(S_t = s|y_t, S_{t-1}). \tag{2.14}$$

Simply, that the probability of transitioning into a state conditioned on the time series up to the point in question is equivalent to the probability of transitioning to a state conditioned on the current time point's value and previous time point's state. The transition probabilities are organized into a transition matrix, denoted usually as $\mathbf{\Pi}$, where the entry π_{ij} denotes $P(S_t = i|S_{t-1} = j)$. In its simplest form the transition matrix is time invariant, however there are modifications that allow for covariates to affect the transition probabilities (Diebold, Lee, & Weinbach 1994). One of the methodological innovations of this dissertation is to modify the transition probabilities as a function of an exogenous variable, but for the explanation of the regime switching GARCH model of Pelletier (2006) we will assume time invariant transition probabilities, that is, the probability of transitioning from one state to the next depends only on values of the observed variables and knowledge of the prior state.

2.1.6 Regime Switching Dynamic Correlation GARCH

The Regime Switching Dynamic Correlation GARCH model of Pelletier (2006) combines univariate GARCH modeling of the variances of each dimension with a regime switching correlation matrix. This reduces the number of parameters to be estimated considerably, and also allows for non-parametric change in the covariance structure. This is ideal for modeling motion artifacts as it permits a flexibility in the effect of motion without adding undue complexity to the model. A pure RSDC-GARCH approach would allow a substantive neuroimager to examine the different correlation regimes one obtained from an fMRI time series, and assess which if any are associated with increased

motion. Later in this dissertation I develop an extension of the RSDC-GARCH, the CTRS-GARCH, that implicitly assesses the association between motion and correlation states.

The RSDC-GARCH model has a similar parameterization as the DCC model (Engle 2002):

$$\mathbf{y}_t \sim N_p(0, \mathbf{D}_t \mathbf{\Gamma}_t \mathbf{D}). \tag{2.15}$$

$$d^2_{(p)t} = \omega_{(p)} + \alpha_{(p)} d^2_{(p)t-1} \epsilon^2_{(p)t-1} + \beta_{(p)} d^2_{(p)t-1}. \tag{2.16}$$

$$\mathbf{\Gamma}_t = \mathbf{\Gamma}_{S_t}. \tag{2.17}$$

Where S_t is the state at time t, follows an unobserved Markov process, and can take values of 1 to K. The RSDC-GARCH model as originally proposed by Pelletier (2006) is very flexible, and the model for $\mathbf{D}_t^2$ can be modified to any type of conditional heteroskedasticity model. Additionally, Pelletier (2006) provides a two step estimation procedure for the entire model, which renders estimation very simple. This estimation procedure is effectively an Expectation-Maximization algorithm that first estimates the univariate GARCH models, and uses those estimates to inform the estimation of the regime shifts in the correlation structure.

2.1.7 The CTRS-GARCH model

The RSDC-GARCH model of Pelletier (2006) provides a way of inferring structural shifts in the covariance matrix of a multivariate time series, and at first glance this is precisely what we need for the estimation of the motion artifact. However, there is an issue with the motion artifact that is not considered.

The issue with using the original RSDC-GARCH model for motion correction is that RSDC-GARCH infers regime changes, and the corresponding covariance structure, without regard to any observed covariates. This means that if there is any sort of dynamic functional connectivity present, the RSDC-GARCH might estimate regimes that are a mixture of the motion artifact and meaningful variation.

As such, any method for motion correction must build in information about the motion itself. Here we will use Frame Displacement (Power, Barnes, et al. 2011), which was discussed in Chapter

1 of this dissertation, as our overall motion covariate. We need to use the Frame Displacement information to inform the regime changes. However, there is one last caveat, and that is that traditionally, regimes are nominal in their ordering, so that regime 2 will not necessarily be more "severe" than regime 1. These unordered regimes with respect to movement would make it more difficult for the researcher to determine which if any of the regimes correspond to a no-motion regime. As we want our regimes to reflect increasingly severe motion artifacts, we need to be able to enforce an ordering on the regimes. In fact, this ordering will simplify estimation considerably when compared to a fully unordered covariate based model described by Diebold et al. (1994). We do this by taking inspiration from the graded response model of Samejima (1969). While this is not strictly necessary for us to estimate the effect of motion on state transition, it will radically simplify interpretation. Specifically, if the states were unordered with respect to motion, conceivably, we could have State 1 be the high motion state, while State 3 is the low or no motion state, and a substantive researcher would have to examine the parameter estimates for the effect of motion to determine which state's correlation matrix they want to use.

The next section has several parts. In the first part, the general form of the CTRS-GARCH model is laid out. The basics of a two step estimation procedure are then presented using a simplified example that does not contain any regimes. This is followed up by the development of the expectation step from an EM algorithm, that will obtain estimates of probability of regime membership given estimates of all other parameters. This exposition will be followed by the specifics of the estimation of the transition matrix, and the parameters governing the relation between motion and regime, this being the innovation of this dissertation. What follows is the primary innovation of this dissertation, a specification of the CTRS-GARCH model.

2.2 The CTRS-GARCH Model Specification

Consider the following RS-GARCH type model:

$$\mathbf{y}_t \sim N_p(0, \mathbf{\Sigma}_t) \qquad \mathbf{\Sigma}_t = \mathbf{D}_t \mathbf{\Gamma}_t \mathbf{D}_t. \tag{2.18}$$

$$d^2_{(p)t} = \omega_{(p)} + \alpha_{(p)} d_{(p)t-1} \epsilon^2_{(p)t-1} + \beta_{(p)} d_{(p)t-1} + \phi_{(p)} x_{t-1}. \tag{2.19}$$

$$\mathbf{\Gamma}_t = \mathbf{\Gamma}_{S_t}. \tag{2.20}$$

Where $\mathbf{y}_t$ is the p dimensional time series, Σ_t is the covariance matrix at time t, D_t is a diagonal matrix, with elements $d_{(p)t}$, each representing a standard deviation, and $\epsilon_{(p)t-1}$ is the normalized residual (if there is a mean structure model), or simply the normalized value (if there is no mean structure, such as in fMRI data) for the pth dimension at $t-1$ (therefore having a variance of 1) . $\omega_{(p)}$ is the intercept for the pth variance, $\alpha_{(p)}$ is the lag-1 parameter for the first GARCH component, while $\beta_{(p)}$ is the second lag-1 parameter for the GARCH expression. Finally, $\phi_{(p)}$ is the effect of x_{t-1} on the pth dimension, where x_{t-1} is the covariate time series. Recall that these dimensions are our ROI time series, and that x_{t-1} can be considered our frame displacement values. This addition of a covariate to the individual variance processes, while not novel to GARCH models as they have been able to handle covariates since development, is a central part of this dissertation. However, this approach to smoothly varying variance correction is novel to fMRI/neuroimaging, particularly with respect to motion.

Γ_t is the correlation matrix at time t, and it is defined as a correlation matrix of regime S_t. We then have to define the transition matrix between these states. The transition matrix between states S is defined with two components. The first is a component that builds in the effect of x, meaning that it is the influence of x_{t-1} on the probability of transitioning into a state at t, regardless of the state one was in at time $t-1$, while the second is a non-time varying Markov component, as traditionally defined, i.e. predicted by the state at the previous time point. These two components are considered independent of each other, and x is considered exogenous to the system, leading to the following

specification for the conditional transition aspect of the CTRS-GARCH[1]:

$$P(S_t = k|x_{t-1}, S_{t-1} = j) \propto \frac{P(S_t = k|x_{t-1})P(S_t = k|S_{t-1} = j)}{P(S_t = k)}. \tag{2.21}$$

Where $P(S_t = k)$ is the prior probability of being in state k. This prior needs to be specified *a priori*, and could potentially be a flat prior, but sensitivity analyses will have to be done.

Going component by component through the above probability in Equation 2.21, $P(S_t = k|x_{t-1})$ can be expressed as a logistic function of x_{t-1}. The unadjusted marginal probability of being in a regime either s or less than s is

$$P^*(S_t \leq s|x_{t-1}) = \frac{\exp(-1.7\zeta(x_{t-1} - \psi_s))}{1 + \exp(-1.7\zeta(x_{t-1} - \psi_s))}. \tag{2.22}$$

Where ψ_s is strictly increasing in s, and can be thought of as the *difficulty parameter*, whereas ζ can be thought of as the slope parameter. 1.7 is a value chosen so that the logistic equation approximates an equivalent probit equation. This is a standard choice stemming from item response theory literature (Samejima 1969).

We adjust it in the following way, following Samejima (1969, Eq. 5)

$$P(S_t = 1|x_{t-1}) = 1 - P^*(S_t = 2|x_{t-1}). \tag{2.23}$$

$$P(S_t = 2|x_{t-1}) = P^*(S_t = 2|x_{t-1}) - P^*(S_t = 3|x_{t-1}). \tag{2.24}$$

$$P(S_t = 3|x_{t-1}) = P^*(S_t = 3|x_{t-1}) - P^*(S_t = 4|x_{t-1}). \tag{2.25}$$

$$\dots \tag{2.26}$$

So on and so forth. This is a cumulative category response function of Samejima (1969), using a logistic model as the basis function. Furthermore, a full explication of $P(S_t = s|x_{t-1})$ is:

[1]This is a central methodological development of this dissertation, and the form comes from the formula to combine two conditional probabilities, if the two conditioning variables are considered different sources of information (i.e. conditionally independent given the outcome): $p(A|D, C) \propto \frac{P(A|D)P(A|C)}{P(A)}$

$$P(S_t = s|x_{t-1}) = \frac{\exp[-1.7\zeta(x_{t-1} - \psi_{s+1})] - \exp[-1.7\zeta(x_{t-1} - \psi_s)]}{[1 + \exp[-1.7\zeta(x_{t-1} - \psi_{s+1})]][1 + \exp[-1.7\zeta(x_{t-1} - \psi_s)]]}. \tag{2.27}$$

The above is directly from Samejima (1969, Eq. 10). Now, we can compute the full combined transition matrix that brings in the static transition matrix, and the covariate information. Let $\mathbf{P}_{X_{t-1}}$ be a diagonal matrix with dimensions $K \times K$, where the ith entry on the diagonal is $P(S_t = i|x_{t-1})$. Let $\mathbf{\Pi}$ be the Markov implied transition matrix. Then the full transition matrix for transition from $t-1$ to t :

$$\mathbf{\Pi}_{x_{t-1}} = \text{diag}((\mathbf{j}'\mathbf{\Pi}\mathbf{P}_{x_{t-1}})^{-1})\mathbf{\Pi}\mathbf{P}_{x_{t-1}}. \tag{2.28}$$

Where $\mathbf{j}$ is a $p \times 1$ vector of 1s[2]. This simply sweeps the covariate based probabilities over the static transition matrix, and then renormalizes appropriately. With these pieces specified, we can layout the full likelihood. The likelihood is as follows:

$$\ell(\mathbf{y}|\boldsymbol{\theta}, \mathbf{x}) = \sum_{t=1}^{T} \log f(\mathbf{y}_t|\mathbf{y}_{t-1}, \boldsymbol{\theta}, x_{t-1}). \tag{2.29}$$

$$\log f(\mathbf{y}_t|\mathbf{y}_{t-1}, \boldsymbol{\theta}, x_{t-1}) = \log((\boldsymbol{\xi}_{t|t-1})'\boldsymbol{\eta}(\mathbf{y}_t)). \tag{2.30}$$

Where

$$\boldsymbol{\xi}_{t|t} = \frac{\boldsymbol{\xi}_{t-1} \odot \boldsymbol{\eta}(\mathbf{y}_t)}{(\boldsymbol{\xi}_{t-1})'\boldsymbol{\eta}(\mathbf{y}_t)}. \tag{2.31}$$

$$\boldsymbol{\xi}_{t|t-1} = \Pi_{x_{t-1}}\boldsymbol{\xi}_{t-1|t-1}. \tag{2.32}$$

$$\boldsymbol{\eta}(\mathbf{y}_t) = \begin{bmatrix} f(\mathbf{y}_t|\mathbf{y}_{t-1}, S_t = 1, \boldsymbol{\theta}, x_{t-1}) \\ \vdots \\ f(\mathbf{y}_t|\mathbf{y}_{t-1}, S_t = K, \boldsymbol{\theta}, x_{t-1}). \end{bmatrix} \tag{2.33}$$

$\boldsymbol{\xi}_{t|t}$ is a p length vector of filtered probabilities of being in each state at time t given all the

[2]Note that this thesis makes use of the diagonal matrix sweep operation. The use of diagonal matrices makes it simple to multiply every row or column of another matrix by a scalar value.

timepoints up to time t, and $\boldsymbol{\xi}_{t|t-1}$ is the probabilities of being in each state at time t given all time points up to time $t-1$. These equations are taken from Pelletier (2006, Eq. 3.1-3.4), which is in turn based on the work of Hamilton (1994). It is worthwhile to note that these two expressions are recursively defined, and need to be estimated using a variation of the Hamilton filter (Hamilton 1989) which has been developed by Diebold et al. (1994).

At this point, it is good to take a step back and determine every parameter we need to estimate. We need to estimate the following:

1. Parameters of the GARCH components.
2. The statewise correlation matrices $\mathbf{\Gamma}_{S_t}$.
3. $\mathbf{\Pi}$, the entries of the static transition matrix.
4. Parameters governing the association of the covariate and the states.
5. Probabilities of regime membership for each time point.

We account for each of these in turn.

2.2.1 GARCH components

In this section, I follow the work of Pelletier (2006), Engle (2002) and Bollerslev (1986) and derive the tools for maximum likelihood estimation of the GARCH components of the overall model. These derivations follow Bollerslev (1986) work, with the trivial exception of adding the $\phi_{(p)}x_{t-1}$ term, and showing that it does not change estimation. Recall that each variance within the covariance structure of our multivariate time series follows a GARCH model that has no regime switching component:

$$d^2_{(p)t} = \omega_{(p)} + \alpha_{(p)} d^2_{(p)t-1}\epsilon^2_{(p)t-1} + \beta_{(p)} d^2_{(p)t-1} + \phi_{(p)}x_{t-1}. \tag{2.34}$$

Our first order of business is to find maximum likelihood estimates for the parameters $\omega_{(p)}$,$\alpha_{(p)}$,$\beta_{(p)}$ and $\phi_{(p)}$. To do this we follow Pelletier (2006)s strategy and separate the likelihood into two more manageable pieces, specifically a likelihood for the GARCH component, and a likelihood for the correlation matrix.

This process is best illustrated with an example. Consider the multivariate normal time series of p dimensions, with:

$$\mathbf{y}_t \sim N_p(0, \mathbf{\Sigma}_t). \tag{2.35}$$

Where $\mathbf{\Sigma}_t$ follows an arbitrary time varying process, possibly the CTRS-GARCH model[3]

The probability density function for any observation y_t is then:

$$f(\mathbf{y}_t|\mathbf{\Sigma}_t) = (2\pi)^{-\frac{p}{2}}|\mathbf{\Sigma}_t|^{-\frac{1}{2}}\exp(-\frac{1}{2}(\mathbf{y}_t)'\mathbf{\Sigma}_t^{-1}\mathbf{y}_t). \tag{2.36}$$

And the log likelihood is:

$$\ell(\mathbf{\Sigma}|\mathbf{y}) = -\frac{1}{2}\sum_{t=1}^{T}(p\log(2\pi) + \log|\mathbf{\Sigma}_t| + (\mathbf{y}_t)'\mathbf{\Sigma}_t^{-1}\mathbf{y}_t). \tag{2.37}$$

We can then apply a decomposition of $\mathbf{\Sigma}_t = \mathbf{D}_t\mathbf{\Gamma}_t\mathbf{D}_t$. Where $\mathbf{D}_t$ is a diagonal matrix which has standard deviations on the diagonal, and $\mathbf{\Gamma}_t$ is a correlation matrix. This decomposition is used extensively by Pelletier (2006) as well as Engle (2002) and others for the same reason we use it here, to make estimation more tractable.

The log-likelihood then becomes:

$$\ell(\mathbf{\Sigma}|\mathbf{y}) = -\frac{1}{2}\sum_{i=1}^{T}(p\log(2\pi) + \log|\mathbf{D}_t\mathbf{\Gamma}_t\mathbf{D}_t| + (\mathbf{y}_t)'(\mathbf{D}_t\mathbf{\Gamma}_t\mathbf{D}_t)^{-1}\mathbf{y}_t). \tag{2.38}$$

We can then separate out terms using determinant and log rules.

$$\ell(\mathbf{\Sigma}|\mathbf{y}) = -\frac{1}{2}\sum_{t=1}^{T}(p\log(2\pi) + 2\log|\mathbf{D}_t| + \log|\mathbf{\Gamma}_t| + (\mathbf{y}_t)'(\mathbf{D}_t\mathbf{\Gamma}_t\mathbf{D}_t)^{-1}\mathbf{y}_t). \tag{2.39}$$

[3]We make this assumption of an arbitrary process to simplify the division of the likelihood into two pieces. If we explicitly work with the CTRS-GARCH likelihood (which is a mixture likelihood), it becomes much more elaborate to separate. However, based on the work of Pelletier (2006) and Engle (2002), this separation can be done for both RSDC-GARCH as well as regular GARCH models, and therefore can be done with CTRS-GARCH.

Adding and subtracting the $(\mathbf{y}_t)'\mathbf{D}_t^{-1}\mathbf{D}_t^{-1}\mathbf{y}_t$ term leads to:

$$\ell(\mathbf{\Sigma}|\mathbf{y}) = -\frac{1}{2}\sum_{t=1}^{T}(\underbrace{p\log(2\pi) + 2\log|\mathbf{D}_t| + (\mathbf{y}_t)'\mathbf{D}_t^{-1}\mathbf{D}_t^{-1}\mathbf{y}_t}_{\text{Components for the variance models}} + \underbrace{\log|\mathbf{\Gamma}_t| - (\mathbf{y}_t)'\mathbf{D}_t^{-1}\mathbf{D}_t^{-1}\mathbf{y}_t + (\mathbf{y}_t)'(\mathbf{D}_t\mathbf{\Gamma}_t\mathbf{D}_t)^{-1}\mathbf{y}_t)}_{\text{Components for the correlation models}}. \quad (2.40)$$

Note that we can now divide this likelihood into two components, the first for the univariate GARCH model, and the second for the correlation matrix. We will deal with the correlation matrix for the CTRS-GARCH model in the next section, so will not be working with that here. Here, we focus only on the univariate GARCH models. The above derivations have shown that one can partition the likelihood into a form that lets one estimate the univariate GARCH models without any concern for the correlation matrix, and be guaranteed consistency of the GARCH models estimation. This follows from the work of Engle (2002).

$$\ell_{\mathbf{D}}(\mathbf{y}) = -\frac{1}{2}\sum_{i=1}^{T}(p\log(2\pi) + \log|\mathbf{D}_t\mathbf{D}_t| + (\mathbf{y}_t)'\mathbf{D}_t^{-1}\mathbf{D}_t^{-1}\mathbf{y}_t. \quad (2.41)$$

As $\mathbf{D}_t$ is a diagonal matrix, this likelihood can be further simplified:

$$\ell_{\mathbf{D}}(\mathbf{y}) = -\frac{1}{2}\sum_{t=1}^{T}\left[p\log(2\pi) + \sum_{j=1}^{p}2\log(d_{(j)t}) + \sum_{j=1}^{p}\frac{y_{(j)t}^2}{d_{(j)t}^2}\right]. \quad (2.42)$$

Some minor rearranging leads us to the exact expression as in Engle (2002, Eq. 30):

$$\ell_{\mathbf{D}}(\mathbf{y}) = -\frac{1}{2}\sum_{t=1}^{T}\sum_{j=1}^{p}\left[\log(2\pi) + \log(d_{(j)t}^2) + \frac{y_{(j)t}^2}{d_{(j)t}^2}\right] \quad (2.43)$$

Note that this expression is a sum of the individual likelihoods of each dimension. This means that each dimension's corresponding GARCH model for the variance can be estimated separately (Engle 2002).

We can now determine the gradient and Hessian for each of the GARCH component parameters.

Recall that our expressions for any dimension's variance is:

$$d^2_{(j)t} = \omega_{(j)} + \alpha_{(j)} d^2_{(j)t-1} \epsilon^2_{(j)t-1} + \beta_{(j)} d^2_{(j)t-1} + \phi_{(j)} x_{t-1}. \tag{2.44}$$

Again, where x_{t-1} is our value of the covariate, in our application Frame Displacement.

The likelihood for any individual dimension's variance is:

$$\ell_{d^2_{(j)}}(x) \propto -\frac{1}{2} \sum_{t=1}^{T} \log(d^2_{(j)t}) + \frac{y^2_{(j)t}}{d^2_{(j)t}}. \tag{2.45}$$

To make notation a bit simpler to follow, I will change $d^2_{(j)t}$ to $h_{(j)t}$ here.

Following the process of Bollerslev's (1986) derivation of the GARCH model maximum likelihood estimates, we can group our $\omega_{(j)}$, $\alpha_{(j)}$, $\beta_{(j)}$ and $\phi_{(j)}$ into a single vector, $\boldsymbol{\theta}_{h_{(j)}}$ and take the derivative of the likelihood with respect to that vector:

$$\frac{\partial \ell_{h_{(j)}}}{\partial \boldsymbol{\theta}_{h_{(j)t}}} = -\frac{1}{2} \sum_{t=1}^{T} h^{-1}_{(j)t} \frac{\partial h_{(j)t}}{\partial \boldsymbol{\theta}_{h_{(j)}}} + \frac{y^2_{t,j}}{h^2_{(j)t}} \frac{\partial h_{(j)t}}{\partial \boldsymbol{\theta}_{h_{(j)}}} \tag{2.46}$$

$$= \frac{1}{2} \sum_{t=1}^{T} h^{-1}_{(j)t} \frac{\partial h_{(j)t}}{\partial \boldsymbol{\theta}_{h_{(j)}}} \left(\frac{y^2_{(j)t}}{h_{(j)t}} - 1 \right). \tag{2.47}$$

Which is identical to Equation 19 in Bollerslev (1986).

Similarly, we can derive the Hessian matrix:

$$\frac{\partial^2 \ell_{h_{(j)t}}}{\partial \boldsymbol{\theta}_{h_{(j)}} \partial \boldsymbol{\theta}'_{h_{(j)}}} = \frac{\partial}{\partial \boldsymbol{\theta}'_{h_{(j)}}} \left(\frac{1}{2} h^{-1}_{(j)t} \frac{\partial h_{(j)t}}{\partial \boldsymbol{\theta}_{h_{(p)}}} \right) \left(\frac{y^2_{(j)t}}{h_{(j)t}} - 1 \right) - \frac{1}{2} h^{-2}_{(j)t} \frac{\partial h_{(p)t}}{\partial \boldsymbol{\theta}_{h_{(j)}}} \frac{\partial h_{(p)t}}{\partial \boldsymbol{\theta}'_{h_{(p)}}} \frac{y^2_{(j)t}}{h_{(j)t}}. \tag{2.48}$$

The two relevant derivatives are now $\frac{\partial h_{(j)t}}{\partial \boldsymbol{\theta}_{h_{(j)}}}$ and $\frac{\partial^2 h_{(j)t}}{\partial \boldsymbol{\theta}^2_{h_{(j)}}}$, which are as follows:

$$\frac{\partial h_{(j)t}}{\partial \boldsymbol{\theta}_{h_{(j)}}} = \begin{bmatrix} 1 \\ \frac{\partial h_{(j)t-1}\epsilon^2_{(j)t-1}}{\partial \alpha_j} \\ \frac{\partial h_{(j)t-1}}{\partial \beta_j} \\ x_{t-1} \end{bmatrix}. \tag{2.49}$$

and

$$\frac{\partial^2 h_{(j)t}}{\partial \boldsymbol{\theta}^2_{h_{(j)}}} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{\partial^2 h_{(j)t-1}\epsilon^2_{(j)t-1}}{\partial \alpha_j^2} & \frac{\partial^2 h_{(j)t-1}\epsilon^2_{(j)t-1}}{\partial \alpha_j \partial \beta_{(j)}} & 0 \\ 0 & \frac{\partial^2 h_{(j)t-1}\epsilon^2_{(p)t-1}}{\partial \partial \beta_{(j)} \alpha_j} & \frac{\partial^2 h_{(j)t-1}\epsilon^2_{(j)t-1}}{\partial \beta^2_{(j)}} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{2.50}$$

As in Bollerslev (1986), the gradient and the Hessian are defined recursively due to the presence of $h_{(p)t-1}$ in the derivatives. Therefore, we repeat the use of the Bollerslev (1986) estimation routine, which he developed from the work of Berndt, Hall, Hall, & Hausman (1974). Simply, we use an iterative approach similar to Fisher Scoring, but instead of using the analytical information matrix, we use a consistent estimator (Bollerslev 1986, pg. 317):

$$\boldsymbol{\theta}^{(i+1)}_{h_{(j)}} = \boldsymbol{\theta}^{(i)}_{h_{(j)}} + \lambda_i \left(\sum_{t=1}^{T} \frac{\partial \ell_{h_{(j)t}}}{\partial \boldsymbol{\theta}_{h_{(j)}}} \frac{\partial \ell_{h_{(j)t}}}{\partial \boldsymbol{\theta}'_{h_{(j)}}} \right)^{-1} \sum_{i=1}^{T} \frac{\partial \ell_{h_{(j)t}}}{\partial \boldsymbol{\theta}_{h_{(j)}}}. \tag{2.51}$$

Where λ_i is an adaptive step size. Again, as in Bollerslev (1986), the maximum likelihood estimate of $\boldsymbol{\theta}_{h_{(j)}}$ under the assumption that the model is correctly specified has an asymptotic distribution with the mean being the true parameter set $\boldsymbol{\theta}_{h_{(j)0}}$ and covariance structure $E(\frac{\partial \ell_{h_{(j)t}}}{\partial \boldsymbol{\theta}_{h_{(j)}}} \frac{\partial \ell_{h_{(j)t}}}{\partial \boldsymbol{\theta}'_{h_{(j)}}})$, which in turn has a consistent sample estimate of $\frac{1}{T} \sum_{i=1}^{T} \frac{\partial \ell_{h_{(j)t}}}{\partial \boldsymbol{\theta}_{h_{(j)t}}} \frac{\partial \ell_{h_{(j)t}}}{\partial \boldsymbol{\theta}'_{h_{(j)}}}$.

We now move onto the estimation of the state specific correlation matrices.

2.2.2 Estimation of State Specific Correlation Matrices

With the variance processes estimated from the previous section, we can evaluate regime specific correlation matrices. Recall from our definition of the CTRS-GARCH model and the decomposition

of the covariance matrix that:

$$\mathbf{y}_t \sim N_p(0, \mathbf{D}_t \mathbf{\Gamma}_{S_t} \mathbf{D}_t). \tag{2.52}$$

Where $\mathbf{\Gamma}_{S_t}$ is the correlation matrix of regime S_t. Fortunately, other authors have already determined that the maximum likelihood estimate of $\mathbf{\Gamma}_{S_t}$ is a simple weighted average of the sample implied correlation matrices per time point. Specifically, Pelletier (2006, Eq. 3.10), citing Hamilton (1994) shows that the maximum likelihood regime specific correlation matrix is:

$$\hat{\mathbf{\Gamma}}_{S_t=s} = \frac{\sum_{t=1}^{T} (\mathbf{u}_t \mathbf{u}_t') \boldsymbol{\xi}_{S_t=s,t|T}}{\sum_{t=1}^{T} \boldsymbol{\xi}_{S_t=s,t|T}}. \tag{2.53}$$

Where $\xi_{S_t=s,t|T}$ is the posterior probability of regime membership of time point t, given the full time series (a smoothed posterior probability, if one will). $\mathbf{u}_t$ is the *estimate* of the standardized vector of $\mathbf{y}_t$, standardized with regards to the estimated variance for timepoint t. So, as one can see, the estimate of the correlation matrix depends in part on the estimate of the univariate GARCH models from the previous section. With that, we must proceed on to the estimation of the elements in the transition matrix $\mathbf{\Pi}_{x_{t-1}}$

2.2.3 Estimation of the Static Transition Matrix $\mathbf{\Pi}$

Recall that the entries in our conditional transition matrix $\mathbf{\Pi}_{x_{t-1}}$ can be expressed as:

$$P(S_t = i | x_{t-1}, S_{t-1} = j) \propto \frac{P(S_t = i | x_{t-1}) P(S_t = i | S_{t-1} = j)}{P(S_t = i)}. \tag{2.54}$$

This specification enforces conditional independence between the Markov mechanics of the regime switching, and the covariate impact of on the next time point's regime. The following applies:

$$P(x_{t-1} | S_t = k) \perp P(S_{t-1} = d | S_t = k). \tag{2.55}$$

What this suggests is that we can calculate each component, the static $P(S_t = k | S_{t-1} = d)$ and the covariate based $P(S_t = k | x_{t-1})$ separately, and then combine the estimates to obtain $\mathbf{\Pi}_{z_{t-1}}$. We will start with the estimation of $P(S_t = k | S_{t-1} = d)$ or the $\mathbf{\Pi}$ matrix.

We rely on the results of Hamilton (1994) and Hamilton (1990), who show that when z is strictly

exogenous, then the expected value of the full transition matrix given a set of regime wise variance parameters $\boldsymbol{\theta}$ is:

$$\hat{\pi}_{i,j} = \frac{\sum_{t=2}^{T} P(S_t = i, S_{t-1} = j | \mathbf{u}_1, \ldots, \mathbf{u}_T, \boldsymbol{\theta})}{\sum_{t=2}^{T} P(S_{t-1} = j | \mathbf{u}_1, \ldots, \mathbf{u}_T, \boldsymbol{\theta})}. \tag{2.56}$$

Here, our goal is not to estimate the completed transition matrix, but rather the static component of it, $\mathbf{\Pi}$. As the two components of the transition probability are conditionally independent given $S_t = s$, this implies that if we marginalize over, in this case, the values of x_t, and do not consider them in our calculation, the resulting single matrix $\hat{\mathbf{\Pi}}$ will be a consistent estimator of $E(\mathbf{\Pi}|\boldsymbol{\theta})$.

2.2.4 Estimation of the effect of x on regime change

Here, I use identical logic to the previous section. If marginalizing over x_{t-1} allows one to obtain an estimate of Π, marginalizing over S_{t-1} would allow one to get estimates of the effect of the covariate x_{t-1} on $S_t = i$. Recall the model for this effect:

$$P^*(S_t \leq i | x_{t-1}) = \frac{\exp(\zeta(z_{t-1} - \psi_i))}{1 + \exp(\zeta(x_{t-1} - \psi_i))}. \tag{2.22}$$

Recall too that we have the smoothed probabilities of regime membership for each time point $\boldsymbol{\xi}_t$. We can transform our smoothed probabilities of individual regime memberships into the form specified by Equation 2.22 by solving for P^*. Note that this is similar in form to Samejima (1969) Graded Response Model, though simpler as we do not have an unobserved latent variable needing estimation. Denote this vector of probabilities as $\boldsymbol{\xi}_t^*$. The model for the effect of x (the covariate) on these regime memberships is then:

$$\boldsymbol{\xi}_{t|T}^* = \begin{bmatrix} \frac{\exp(-1.7\zeta(x_{t-1} - \psi_1))}{1+\exp(-1.7\zeta(x_{t-1} - \psi_1))} \\ \vdots \\ \frac{\exp(-1.7\zeta(x_{t-1} - \psi_K))}{1+\exp(-1.7\zeta(x_{t-1} - \psi_K))} \end{bmatrix}. \tag{2.57}$$

Apply an inverse logit to both sides:

$$logit^{-1}(\boldsymbol{\xi}^*_{t|T}) = \begin{bmatrix} \zeta(x_{t-1} - \psi_1) \\ \vdots \\ \zeta(x_{t-1} - \psi_K) \end{bmatrix}. \tag{2.58}$$

As this is an example of a generalized linear method and would not have any values in $\boldsymbol{\xi}_{t|T}$ equal to 0 or 1, one can find the maximum likelihood estimates of β and θ_S using ordinary least squares. An easy procedure is as follows.

First stack $logit^{-1}(\boldsymbol{\xi}^*_t)$ so that the resulting vector $\mathbf{a}$ is $((T * K) \times 1)$. Next construct $\delta_{i,S}$ such that $\delta_{i,S} = 1$ if a_i was originally from an element of $logit^{-1}(\boldsymbol{\xi}^*_t)$ corresponding to $S_t = S$. $\mathbf{a}$ can then be modeled as $a_i = \zeta(x_{(i)} - \sum_{j=1}^{K} \delta_{i,j}\psi_j)$, and this becomes an exercise in numerical optimization.

2.2.5 Estimation of posterior probabilities of regime membership

At this point we have estimates of the univariate GARCH models, the regime specific correlation matrices, and the conditional transition matrix. Of course, these estimates are all given particular values of the posterior probabilities of regime membership for each point. The key now is to obtain the expected values of the posterior probabilities of regime membership for each time point, given the rest of the parameters in our model.

Fortunately, this step is easy, and Diebold et al. (1994) developed an algorithm for obtaining $\boldsymbol{\xi}_{t|T}$. The algorithm itself is quite extensive, and will not be presented here in full in this dissertation. In essence, one must first calculate the filtered probabilities of regime membership $P(S_t, S_{t-1}|\mathbf{y}_t, x_{t-1}, \boldsymbol{\theta})$, where $\boldsymbol{\theta}$ denotes our general set of parameters for all other elements in the model. Using this set of filtered probabilities, which exist for each time point, we calculate the smoothed joint probabilities of regime membership for all sequential time points, which are then summed over all of the s_{t-1} elements to finally obtain $P(S_t = i|\mathbf{Y}_T, \mathbf{X}_T, \boldsymbol{\theta})$.

2.2.6 The EM Algorithm

Now that we have ways of estimating each one of our parameters, an overall algorithm can be developed. Denote the vector of parameters for the regime specific correlation matrices, the static transition matrix and the effect of the covariate on the state as $\boldsymbol{\psi}$. Denote the parameters for the

univariate GARCH models as $\boldsymbol{\theta}_G$. Denote the smoothed posterior probabilities of class membership as $\boldsymbol{\xi}_{t|T}$. We can then use an EM algorithm (Dempster, Laird, & Rubin 1977), as follows:

1. Obtain maximum likelihood estimates of $\boldsymbol{\theta}_G$ using the procedure from section 2.1.6.
2. Initialize algorithm with starting values $\boldsymbol{\psi}_0$
3. $i = 0$
4. Compute $E[\boldsymbol{\xi}_{t|T}|\boldsymbol{\psi}_i]$ using the Diebold et al. (1994) algorithm.
5. Compute $\boldsymbol{\psi}_{i+1}$ using the algorithms laid out in previous sections.
6. i = i + 1
7. Repeat steps 4 through 6 until $\boldsymbol{\psi}_{i+1}$ has converged.

Under minor regularity condition, the EM algorithm will converge to maximum likelihood estimates of $\boldsymbol{\theta}$ (Dempster et al. 1977).

3 SIMULATIONS

The performance of the CTRS-GARCH is evaluated with two sets of simulations, and compared against the most popular motion correction methods, scrubbing (the wholesale removal of points that exceed a certain motion threshold as indicated by Frame Displacement and DVARS) and 24 parameter motion regression.

There will be two sets of conditions corresponding to two possible types of motion artifacts. The first type is that of discrete motion artifacts, where effect of motion follows a CTRS-GARCH model. This simulation set allows us to both evaluate the performance of our model, and to compare the performance of the CTRS-GARCH to that of scrubbing under optimal conditions for the CTRS-GARCH model. The second set of simulations are where the motion artifact follows a smoothly varying model, which would be difficult to estimate on real data, but easy enough to simulate from.

3.0.1 Simulation of Motion

Of key interest here is the method by which we simulate motion in our design. Following the SimTB Matlab package (Erhardt, Allen, Wei, Eichele, & Calhoun 2012), which is a widely used software package for the simulation of fMRI data, the full 6 parameter of motion, translation in the x, y, z planes, and the three directions of rotation are simulated using AR(1) processes of the following nature:

$$M_{t+1} = .95M_t + \epsilon. \tag{3.1}$$

Where M is the given motion parameter, the autoregressive effect is .95, and ϵ is a random error term with distribution $N(0, \sigma_M)$. These parameters are chosen to be similar to those used in motion parameter simulations of the SimTB Matlab package (Erhardt et al. 2012). However, in empirical data, motion tends to demonstrate its own regime structure, where isolated periods of more motion

alternate with periods of relative stillness. A simple AR(1) process results in a stationary FD sequence, which corresponds to an individual continually making small motions throughout the scan. To reflect the regime switching nature of motion in empirical data, we use a simple regime switching model, with three regimes corresponding to little motion, medium motion and high motion. The transition matrix is as follows:

$$\begin{bmatrix} .8 & .2 & 0 \\ .2 & .6 & .2 \\ .1 & .2 & .7 \end{bmatrix} \tag{3.2}$$

Where this is a matrix of probabilities of transition from one regime to another. This matrix is read row to column, where, for example, row 1, column 2 corresponds to the probability of transitioning from the no motion to the medium motion regime. This matrix was chosen through experimentation to allow for a moderate to large amount of motion, that also has sustained sections of no movement.

These regimes then correspond to the standard deviation of the error term in the AR1 process, ϵ_M. The no motion regime has a corresponding standard deviation of .02, medium motion .05 and high motion .1. These values were chosen so that the resulting frame displacement values were scaled similarly to empirical data, specifically so that the mean frame displacement is around .4, which represents a higher motion subject. FD is then calculated in accordance to Eq 1.3. Figure 3.1 below shows an example time series of the generated frame displacement values, while Figure 3.2 shows an example of a real subject's frame displacement. The simulated data contains considerably more motion than the real data, however both exhibit a clear regime structure to motion (evidenced by sustained spikes in frame displacement.)

The generation of a Frame Displacement time series is the first step of our simulation procedure. We follow this with a construction of the true signal, and the motion contaminated signal.

3.0.2 Generating the True Signal

Here, the true signal represents the BOLD signal uncontaminated by motion. In order to simulate a time series that represents the true signal, we need a true signal correlation matrix. To generate this we follow the lead of Smith et al. (2011), and propose a directed and weighted causal network of the

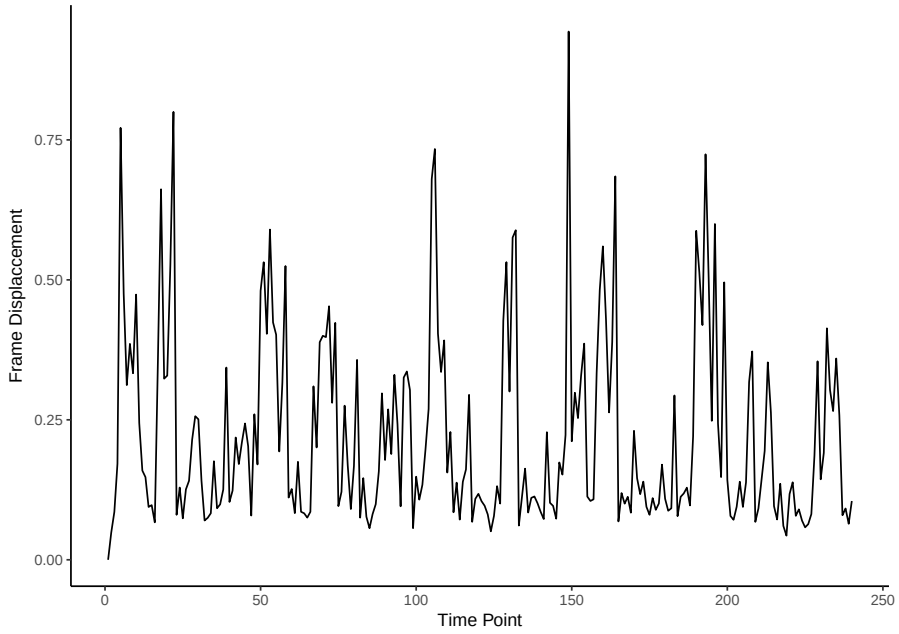


Figure 3.1: Example Simulated Frame Displacement

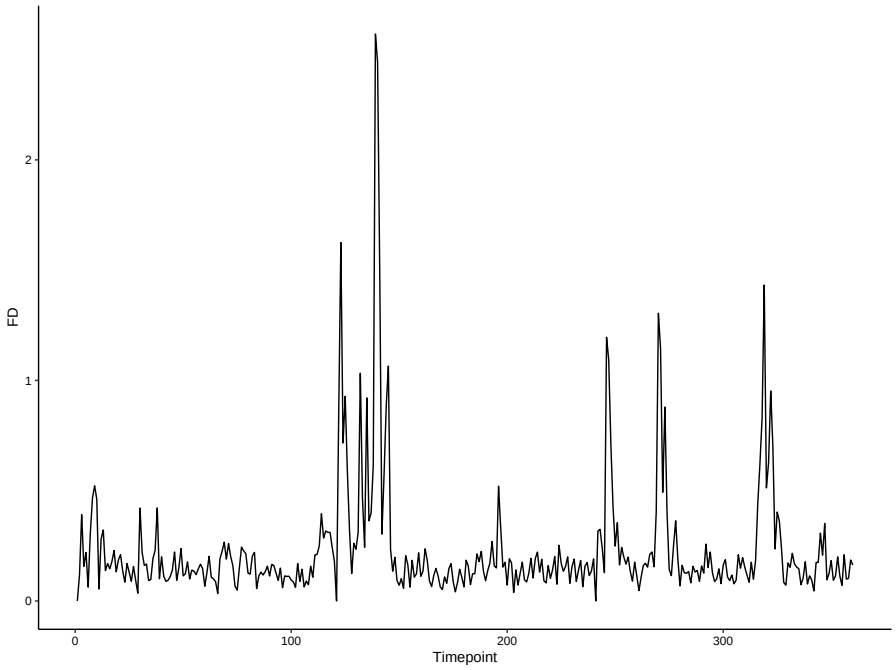


Figure 3.2: Example Empirical Frame Displacement

relations between our synthetic ROIs. Like in the Smith et al. (2011) simulations, the connectivity networks exhibit a community structure (corresponding to the existence of functional networks within the brain), with between community connections. We created four connectivity networks in total

corresponding to four generating conditions, 10 and 15 ROIs, and high and low strength of the signal[1].

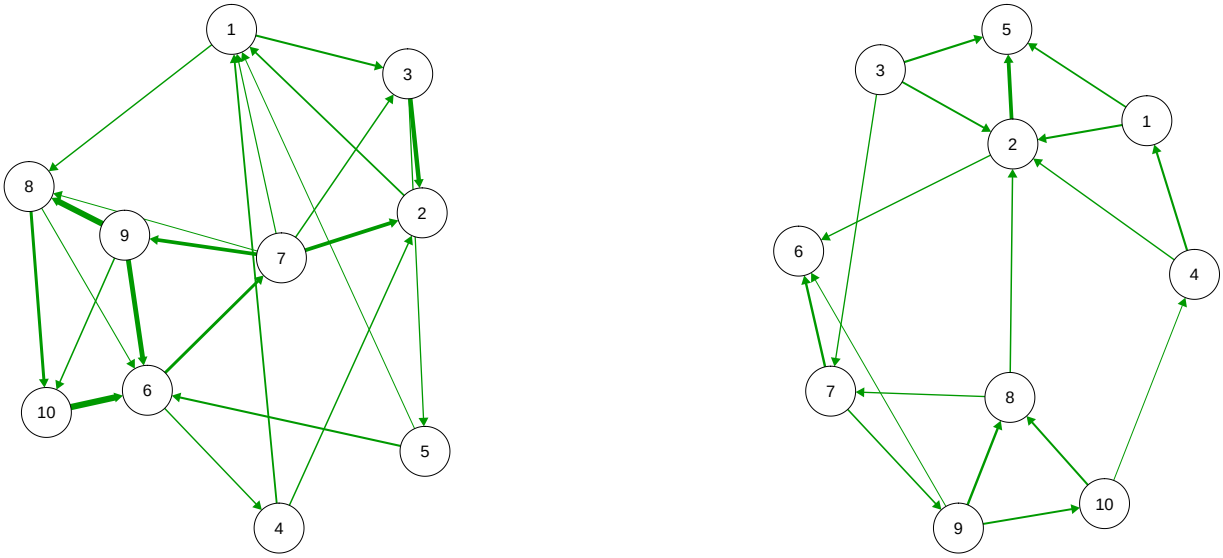
The steps to create these weighted networks are as follows:

1. Generate a vector of community labels with equal probability to belong to each community. 2 communities for the 10 ROI condition, 3 communities for the 15 ROI condition.

2. For each possible edge in a directed network between the ROIs, toggle an edge with probability .2 if the edge is between communities, and .6 if the edge is within a community.

3. For each existent edge assign a weight randomly sampled from $|N(\mu, 1)|$, here $\mu = 1$ in the high signal condition, and $.5$ in the low signal position.

Plots of each network are presented in Figures 3.3 and 3.4.



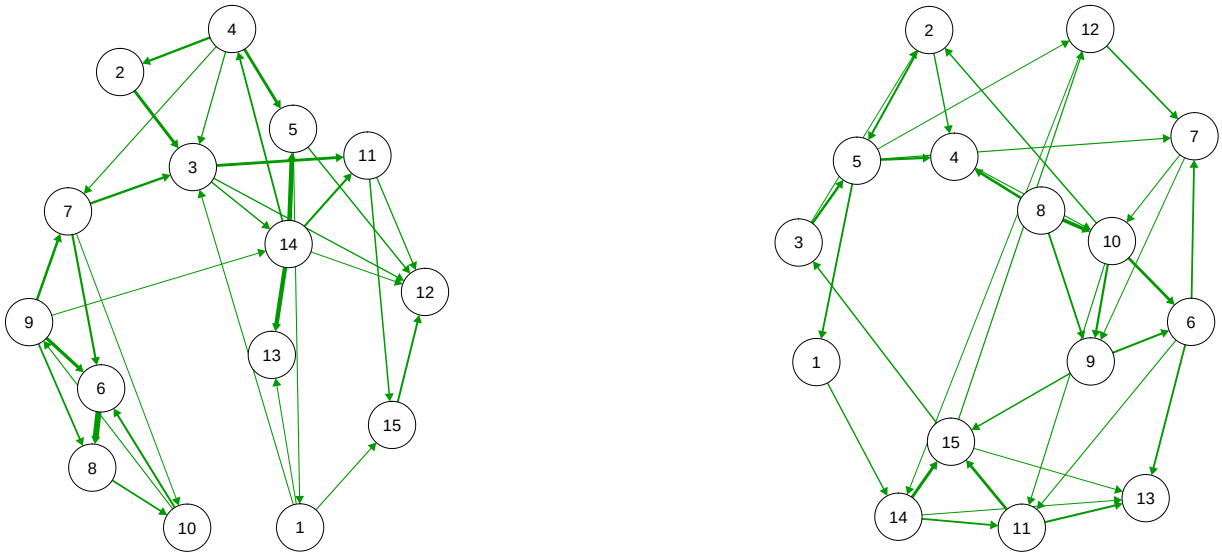
(a) High Signal

(b) Low Signal

Figure 3.3: 10 ROI connectivity networks. Thicker lines represent higher edge weights.

As functional connectivity analysis typically uses the correlation matrix of an fMRI time series, we convert the above functional connectivity networks, which represent the data generating process

[1]the term signal is often used in describing the results and needs to be defined. In this dissertation we using *true signal* to mean the motion free simulated fMRI signal. High and low signal simply means that the high signal conditions were simulated so that there are stronger true correlations over all, while the low signal condition was simulated to have lower true correlations over all.



(a) High Signal

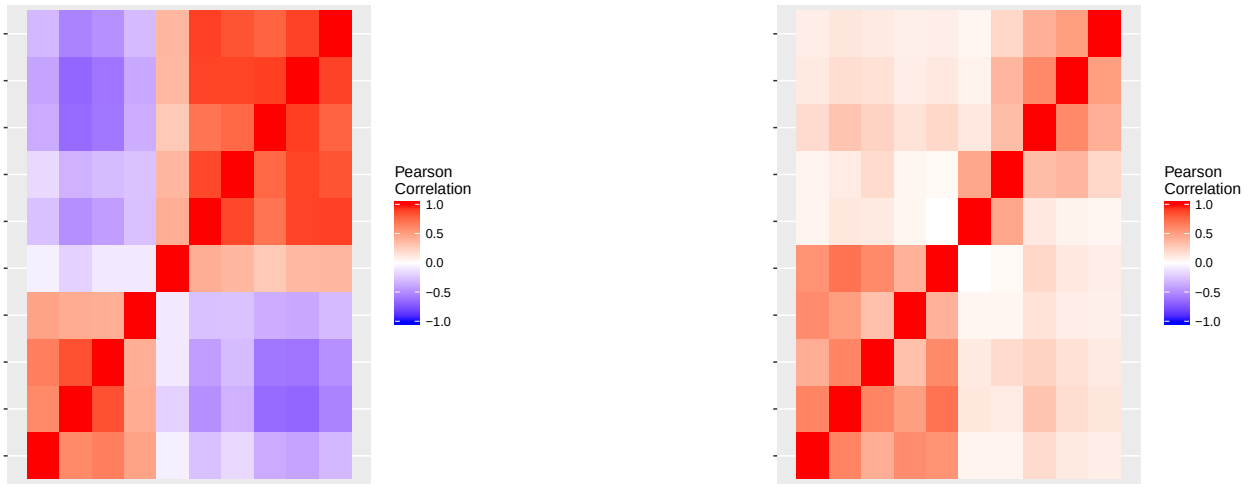
(b) Low Signal

Figure 3.4: 15 ROI connectivity networks. Thicker lines represent higher edge weights.

on the level of the BOLD signal, to correlation matrices. This is done by treating the network as a square matrix of regression coefficients, and using the model implied covariance matrix equation from Bollen (1989):

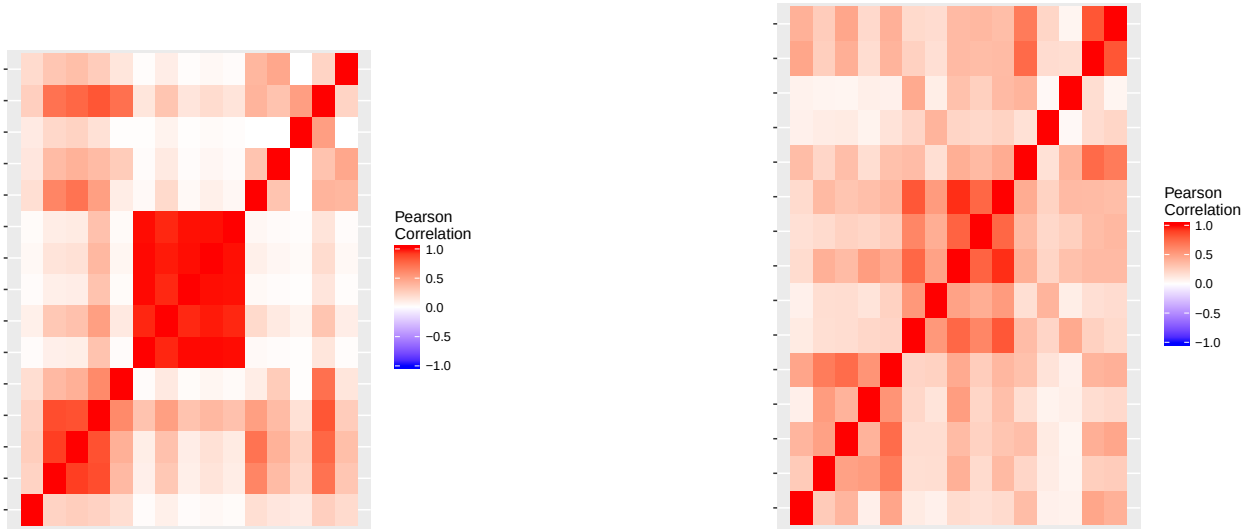
$$\Sigma = (I - B)^{-1}(I - B)^{-1^T}. \tag{3.3}$$

Where Σ is the implied covariance matrix, I is an identity matrix, and B is the matrix representing the connectivity matrix. Σ is then converted into a correlation matrix through the standard algebra. The correlation matrices from each implied network are shown below in Figures 3.5 and 3.6.



(a) High Signal (b) Low Signal

Figure 3.5: 10 ROI Correlation Matrices.



(a) High Signal (b) Low Signal

Figure 3.6: 15 ROI Correlation Matrices.

3.0.3 The Simulated Effect of Motion

With the true signal correlation matrices now generated, the effect of motion needs to be defined. Building off of the work of Power, Cohen, et al. (2011a) and Power et al. (2014), we note that motion tends to increase the correlation between spatially close ROIs, and attenuates the correlation between spatially distant ROIs. To replicate this in our simulations, we use a Toeplitz correlation matrix. An example of this matrix is presented in Figure 3.7. A Toeplitz matrix is defined as follows:

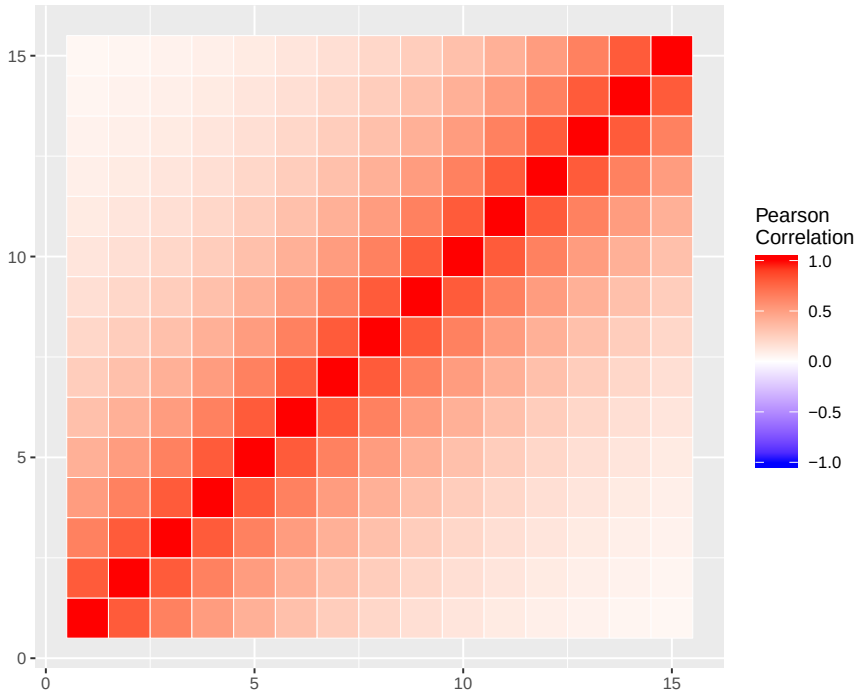


Figure 3.7: Simulated Motion Correlation Matrix

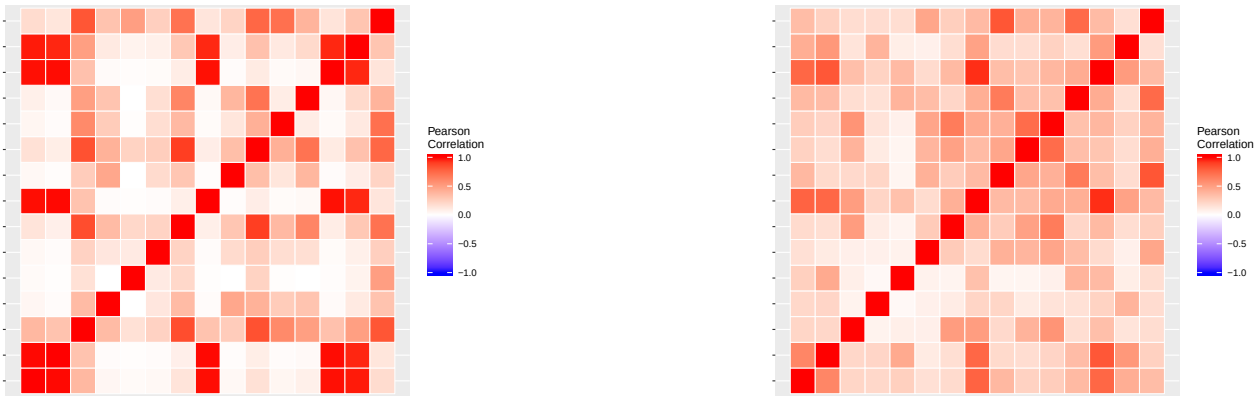
- Let ρ be the target correlation, for example .8.
- The entry ij of the Toeplitz matrix is equal to $\rho^{|i-j|}$.
- This results in a diagonal of 1, and a progressively decreasing correlation as one moves away from the diagonal.
- Toeplitz matrices defined this way are positive definite, and therefore proper correlation matrices.

This Toeplitz matrix can can be thought of as being organized spatially, so that ROI 1 and ROI 2 are spatially close, and ROI 1 and ROI 15 would be the most spatially distal.

3.0.4 Scrambling the True Signal Matrices

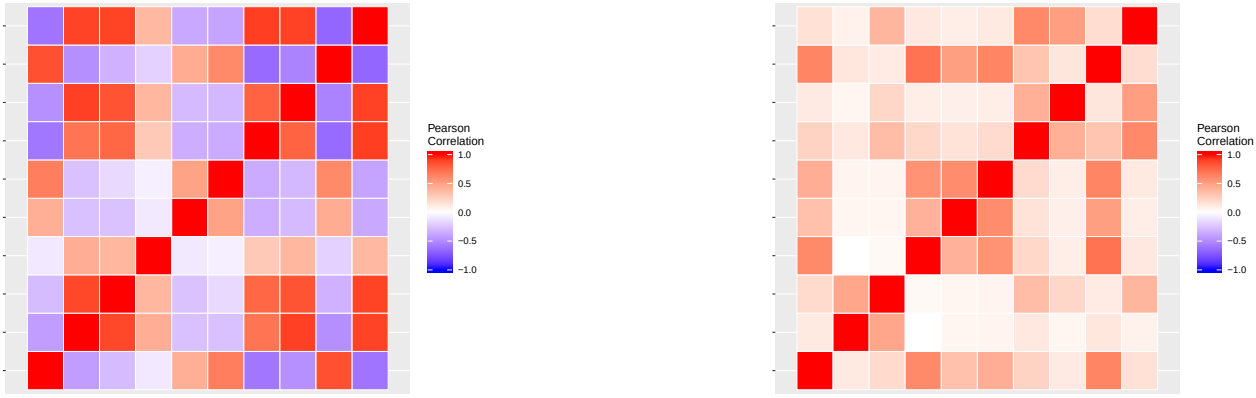
As originally simulated, the correlation matrices for the true signal are in block diagonal form. This would map onto empirical data where clusters were identified *a priori.* As combining the motion and block diagonal matrices would in many ways preserve the block diagonal structure (therefore making it easier to detect the true signal), we scramble the true signal matrices in a deterministic

fashion. This is akin to changing the spatial ordering of the ROIs, so that members of the same functional network are not necessarily spatially close to each other. The "scrambled" correlation matrices are presented in Figures 3.9 and 3.8, and represent the ROIs being spatially ordered so that ROIs close to each other in the correlation matrix are thought to be spatially closer to each other, and vice versa.



(a) High Signal (b) Low Signal

Figure 3.8: 15 ROI Scrambled Correlation Matrices.



(a) High Signal (b) Low Signal

Figure 3.9: 10 ROI Scrambled Correlation Matrices.

3.0.5 Combining True Signal and Motion Correlation Matrices

To create motion contaminated correlation matrices, we take advantage of the fact that the addition of positive definite matrices is itself positive definite. We create motion contaminated matrices using the following equation:

$$\Sigma_{Contaminated} = p\Sigma_{True} + (1-p)\Sigma_{Motion}. \tag{3.4}$$

Where p is a scalar between 0 and 1. By changing the value of p, we can change the degree to which motion has contaminated the correlation matrix. For example, Figure 3.10 shows the high signal 15 ROI matrix contaminated with motion at 70%.

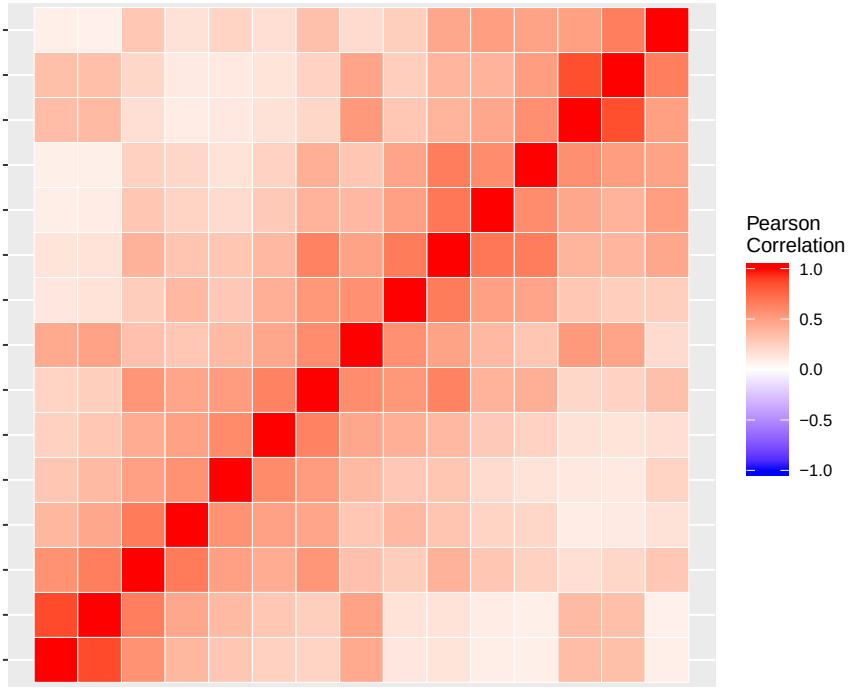


Figure 3.10: 15 ROI High Signal Matrix contaminated at 70% Motion ($\rho = .8$)

3.0.6 Simulating the fMRI time series

The process for simulating the fMRI time series is the same regardless of simulation condition, as each simulation condition simply changes how the aforementioned p, the proportion of motion to true signal, is calculated.

To construct the fMRI time series we use the following algorithm.

1. For each ROI we generate a univariate time series using a GARCH model with the intercept for variance at 1, the lag-1 ARCH parameter at .2 and the lag-1 GARCH parameter at .1, and effect of frame displacement at .5. This creates a set of univariate time series that have smoothly varying variances over the timecourse, but are independent from one another.

2. For each time point, we compute the motion contaminated signal correlation matrix using the simulation condition appropriate method (discrete or continuous, Simulation conditions are described in Section 3.1 below). This creates a set of correlation matrices, one for each time point.

3. For each time point, we take the set of values from that time point across all ROIs (e.g. for 15 ROIs, this would be a set of 15 values), and multiply this vector by the upper-triangular Cholesky decomposition component of the motion contaminated signal correlation matrix. This induces the correct correlation structure between the ROIs.

4. At the end of this procedure we have a multivariate time series where each component has a smoothly varying variance that is partially dependent on motion, and has a correlation structure that changes appropriately with respect to motion. This time series replicates the multivariate time series of an fMRI BOLD signal.

3.1 Simulation Conditions

Both of the following simulation conditions concern the way in which motion contaminates the true signal matrix. In the discrete motion artifacts condition, the effect of motion happens in discrete shifts. If a time point is in the medium motion regime, then the true signal matrix is contaminated at $p = .5$ for example. In the continuous motion artifact condition, the degree of contamination is calculated on the basis of the value of frame displacement at a given time point. For example, a time point with a frame displacement of .1 will be less motion contaminated than a time point with a frame displacement of .5. We now present the details of each condition.

3.1.1 Simulation Set 1: Discrete Motion Artifacts

Recall that motion was generated with a regime structure. In this condition, the regime membership of each timepoint with regards to motion is used to calculate the motion artifact as well. This ensures that the degree of motion contamination both has a regime structure, and those regimes are directly related to observed motion.

We have two conditions within the discrete motion artifacts simulation set. These conditions are the number of regimes for the motion effect (but not motion itself). We use 2 and 3 regimes as

our conditions. The 3 regime condition uses the 3 regimes from the motion generation, while the 2 regime condition collapses the medium and high motion regimes into one regime for the purposes of calculating the motion artifact.

The p values for the 2 and 3 regime condition are 0, .9, and 0, .5, .9, respectively. These values represent a no-motion regime (with $p = 0$), a high motion regime ($p = .9$), and a medium motion regime ($p = .5$). Therefore, the 2 regime condition is technically a harsher condition, as there is no intermediary motion contamination between high contamination and no contamination (that being said, it is also a clearer division, which should make it easier to determine a solution.).

3.1.2 Simulation Set 2: Continuous Motion Artifact

This simulation set operationalizes the effect of motion as a continuous function of frame displacement, rather than a set of motion contamination regimes. To begin, we find the maximum frame displacement of a time series. We then normalize all other values as percentages of the maximum, thus the maximum frame displacement for a time series would have the value 1, half the maximum would have the value .5, etc. We then vary the maximum p into two conditions, where the maximum p is .4, or .8. Thus, the time point with the maximum amount of frame displacement would be motion contaminated at 40% or 80%, depending on which simulation condition the iteration is in. This allows for continuous variation with regards to the level of motion contamination.

3.1.3 Overview of the Simulation Setup.

In total, we have 3 general simulation factors, each with 2 conditions. These are: Number of ROIs (10, 15), number of time points (120, 240), and strength of signal (high and low). These factors are fully crossed for a total of 8 general simulation cells. Within each motion artifact generation method, there is 1 factor with 2 conditions. Within the discrete motion artifact simulation set, this factor is number of regimes (2, 3), and within the continuous motion artifact simulation set, this factor is the maximum amount of motion contamination (.4 and .8). Within a motion artifact generation framework, these conditions are crossed with the general simulation conditions, for a total of 16 cells per motion artifact generation framework, and 32 simulation cells in total. Within each cell, we generate 100 independent samples. Figures 3.11 shows a flowchart representing the process by which data is simulated.

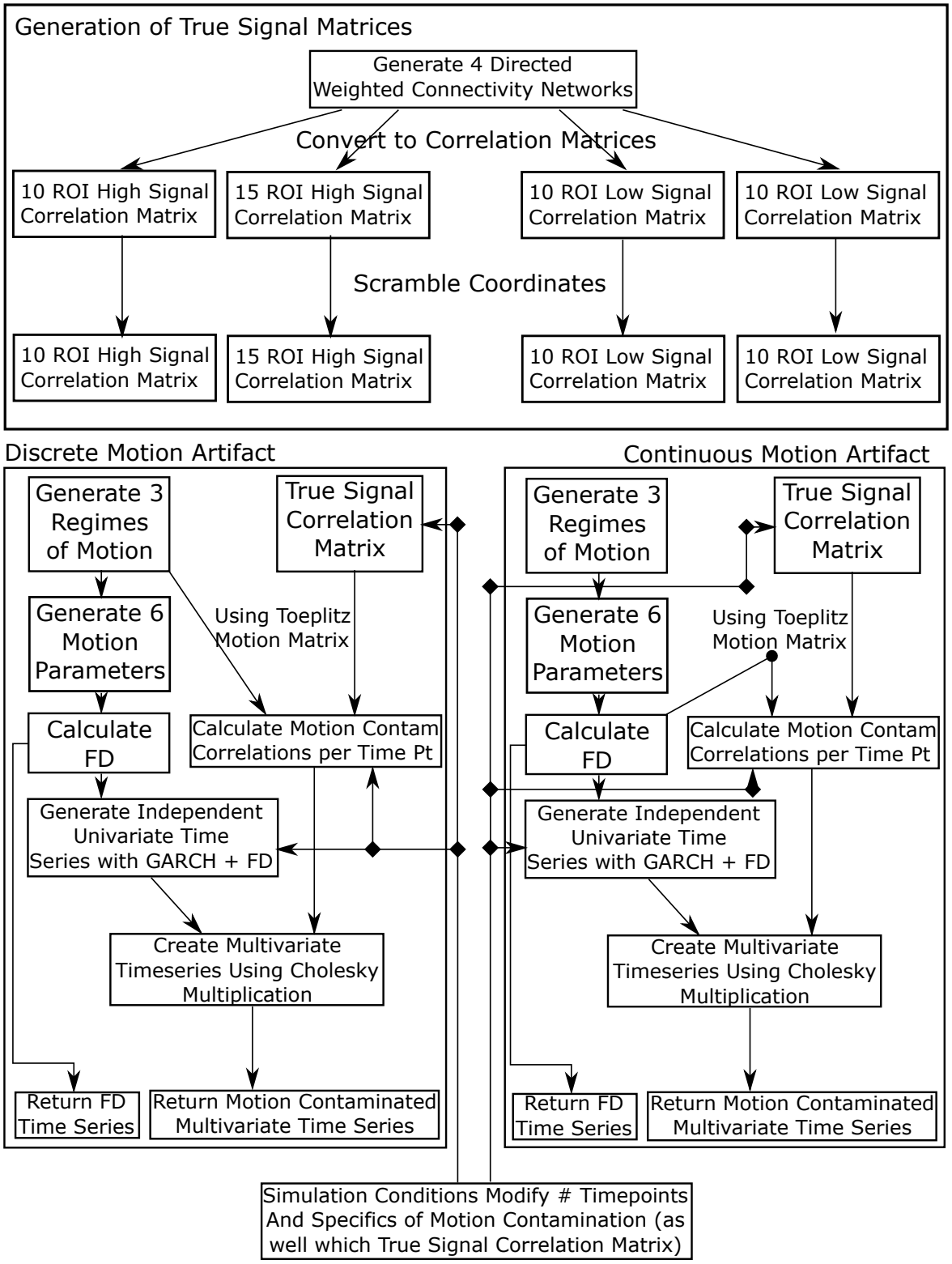


Figure 3.11: A simple visual depiction of the simulation structure. The lower windows are per dataset. Within a set of conditions, this would be repeated 100 times.

3.1.4 Analysis of the Simulated Data

For each of the 100 samples from each of the 32 simulation cells, we apply the CTRS-GARCH method with 2 and 3 classes estimated. Additionally, to compare to the scrubbing approach of Power, Cohen, et al. (2011a), we apply scrubbing at a threshold of .2, .3, .4 and .5 frame displacement. Finally, we use the Friston 24 parameter regression approach, utilizing the 6 motion parameters generated.

Of primary interest to this study is the recovery of the true signal correlation matrix. Each use of CTRS-GARCH produces regime specific correlation matrices, and theoretically, the first regime should be the closest to the true correlation matrix. With the application of scrubbing, and regression, correlation matrices are calculated from the scrubbed time series and the residual time series respectively.

To obtain a one number summary of the quality of the estimated correlation matrices, we use *absolute distance* of the estimated correlation matrix from the true correlation matrix. Additionally, to provide a reference point for this quantity, we compute the *absolute distance* of the full sample (uncorrected) correlation matrix from the true correlation matrix. As correlation matrices have a standard metric, this absolute distance quantity is fully comparable across simulation conditions. Within a simulation condition, we can examine the *improvement* of a motion correction method over the uncorrected data as a decrease in the absolute distance to the true correlation matrix. Furthermore, within a condition we can compare methods by comparing this improvement. We will examine the difference between motion correction methods both qualitatively, as well as quantitatively with paired-samples t-tests. As this is a fairly simple simulation design, larger scale meta-models will not be required to assess differences between condition sets.

3.2 Results

We will present the results[2] from each motion artifact generation condition in turn, and follow that up with a discussion of the general factors (such as time points)

3.2.1 Discrete Motion Artifact

Figures 3.12 and 3.13 show the change in the absolute distance from the true signal correlation, relative to the estimate of the true signal correlation matrix by uncorrected data (the raw data with no motion correction applied to it). This can be interpreted as an improvement, if the change is below 0, or a worsening, if the change is above 0. For example, if the uncorrected estimate is 40 units away from the true correlation matrix, and the correlation matrix produced by CTRS-GARCH was 20 units away, CTRS-GARCH would have produced a better estimate of the true signal correlation matrix than the uncorrected data. However, the values 30 and 20 are relative to the simulation trial that produced them, so to allow our values to be comparable, we subtract the uncorrected data estimate distance from the motion corrected estimate distance. This metric is what is presented in the following figures, and has a simple interpretation. Negative values of this quantity mean that the motion correction method in question performed better than the uncorrected raw data estimate of the correlation matrix, while positive values mean that the motion correction method returned worse values than uncorrected data. Finally, in all sections, we only used the *low-motion* regime of the CTRS-GARCH models for comparison, as all other regimes are expected to be far from the true signal.

[2]A brief note on the reporting of numerical information. As is typical for a simulation study, the following reporting focuses on the degree to which the inferential method recovers the data generating model. This degree of recovery here is the absolute distance between two correlation matrices. While this is an important metric to demonstrate overall performance, its precise numerical value is of considerable less interest. As such in the main text of this dissertation, we refrain from discussing numerical values unless these values provide additional information as to the overall pattern of results.

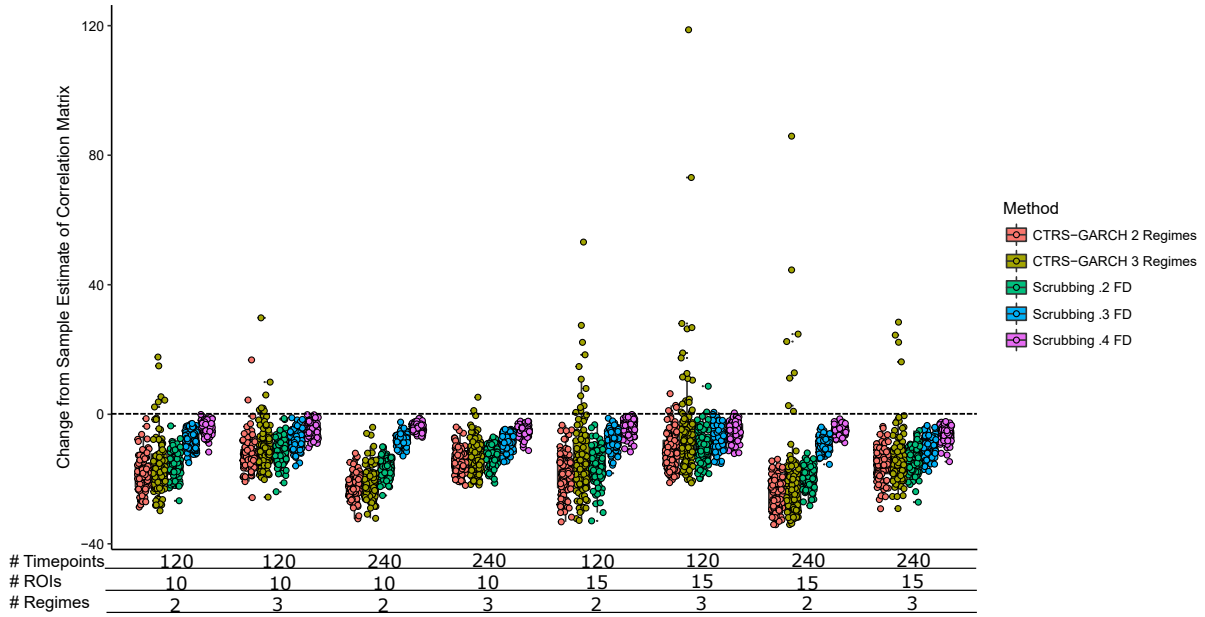


Figure 3.12: Change in the estimate of the true signal correlation matrix for the High Signal Discrete Motion Artifact Condition Set. Note that on average, all presented methods improve the estimate of the true signal correlation matrix, as the majority of the points are below the 0 line. In conditions with more information (more time points, fewer ROIs), 2 regime CTRS-GARCH tends to outperform all other methods. Finally, 3 regime CTRS-GARCH appears to have a tendency to produce outliers, as evidenced by datapoints with extremely bad estimates of the true signal correlation matrix.

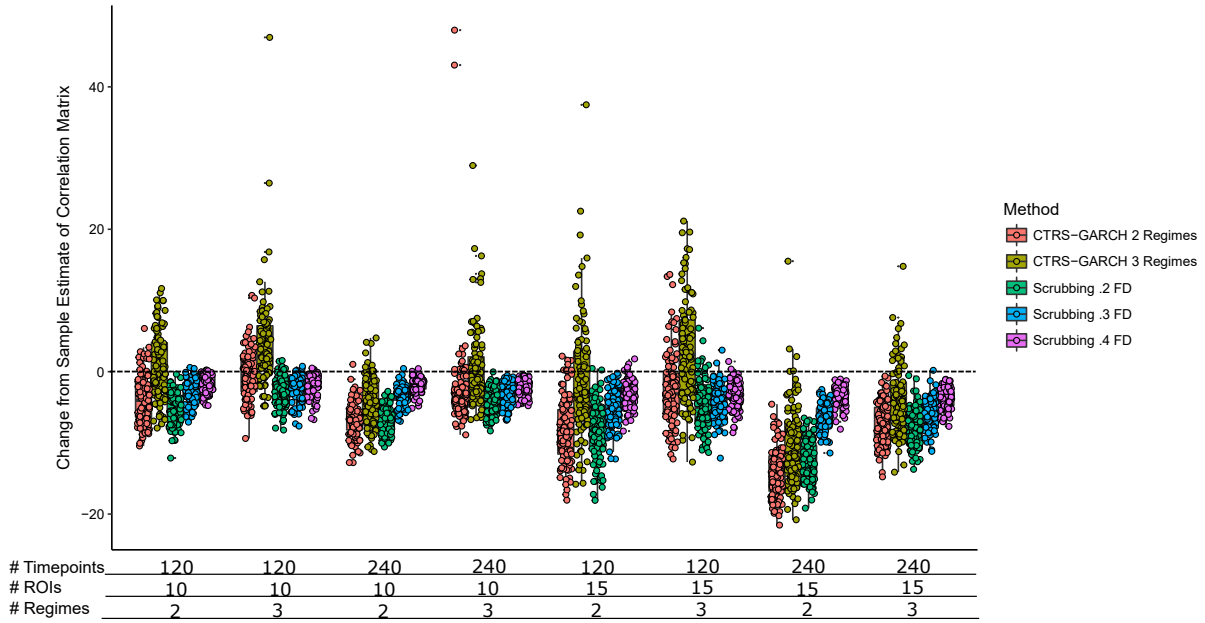


Figure 3.13: Change in the estimate of the true signal correlation matrix for the Low Signal Discrete Motion Artifact Condition Set. This condition set indicates considerably more variance in the relative improvement, with more iterations that resulted in poor estimates of the true signal correlation , particularly for the CTRS-GARCH approaches. These issues appear to be most prevalent at low information conditions, and at higher information conditions, 2 regime CTRS-GARCH tends to either meet or outperform the most conservative scrubbing approach (.2 FD).

The results for the discrete motion artifact condition showed, for the most part, an expected pattern of results. Specifically, for those conditions with a large number of time points, and a stronger true signal correlation, CTRS-GARCH outperformed all other methods[3]. Additionally, a more conservative scrubbing threshold tended to lead to better estimates of the true signal correlation matrix, however, these tended to be slightly more variable. Again, this is to be expected, as a more conservative scrubbing threshold removes more time points, which results in a more variable estimate of the true signal correlation matrix.

Overall, it appears that having a stronger true signal (the High Signal condition), improves the

[3]In all cases across all conditions the regression approach worsened the estimation of the true signal correlation matrix. However, this is not an indictment of the Friston regression approach, but rather this result is likely due to a mismatch between our data simulation method, and the regression approach. Specifically, our data simulation method uses frame displacement as the "cause" of the motion artifact. The regression approach uses the 6 motion parameters as predictors. While the true effect of motion is related directly to these 6 motion parameters, our simulated data is not, so it is not surprising that the regression approach failed here. As such, we do not present the results from the regression approach.

performance of all the motion correction methods reported, both in terms of mean improvement, but also in terms of the variance of the improvements within a condition. This is to be expected, as a stronger signal will have less of a tendency to be washed out by motion (at least in this simulation setup.)

It is, of course, the unexpected results that are most of interest here. To begin, one of the more troubling results is the ability of the CTRS-GARCH method to result in substantially worse estimates of the true signal correlation matrix then the raw uncorrected data would provide. This occurs primarily in conditions of low information (i.e. lower sample size, more ROIs, low true signal strength), and is substantially exacerbated when a 3 regime solution is estimated. The markedly increased variance in the results from a 3 regime CTRS-GARCH is likely due to the simulation structure. While in the three regime simulation conditions, there are three levels of motion contamination, really there are three different levels of mixing of two different distributions, the true signal and the Toeplitz motion contamination matrix. This also provides an explanation for the consistent good performance of the 2 regime CTRS-GARCH solution relative to the 3 regime solution.

Here it is also important to note that while CTRS-GARCH solutions can, potentially, and particularly in low signal conditions, be worse than the raw data, the solutions from scrubbing are, at least in these simulations, very rarely worse than the raw data. This however must be taken with caution, as while the most conservative threshold for scrubbing tended to result in decent estimates of the true correlation matrix, this can also result in a large amount of data loss.

It appears that from these marginal results, CTRS-GARCH is best used with 2 regimes, in conditions where one has a large number of time points, and perhaps a theoretical reason why the true underlying functional signal would be strong. However, there is an additional layer of analysis that needs to be done. As these are relative improvements to the raw data estimated matrix, we can compare within a trial the performance of CTRS-GARCH to that of scrubbing. This is what we present below.

Within iteration comparisons of CTRS-GARCH and Scrubbing

From examination of the marginal plots previously, we can examine the performance of a 2 regime CTRS-GARCH model as compared to scrubbing at the most conservative level (.2 FD), and a more

reasonable level for practice (.3 FD). These plots are presented below in Figures 3.14 and 3.15.

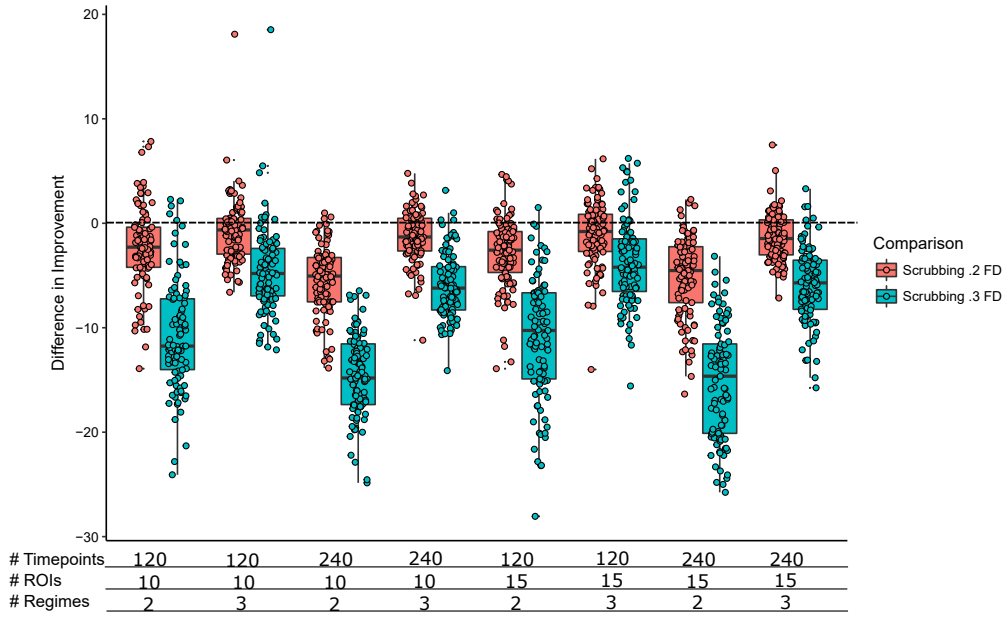


Figure 3.14: Comparison of the 2 regime CTRS-GARCH solution to Scrubbing at .2 and .3 FD within a simulation trial for High Signal Discrete Motion Artifact Conditions. Negative numbers indicate that the CTRS-GARCH solution was superior to the scrubbing solution. Note that on average, 2 regime CTRS-GARCH outperformed scrubbing at both .2 and .3 FD, with the majority of trials indicating that the CTRS-GARCH solution was superior. Again, in higher information conditions, CTRS-GARCH showed clear advantage, whereas for lower information conditions, CTRS-GARCH showed less of an advantage. Note too the large shift in results when looking at .2 FD vs .3 FD scrubbing.

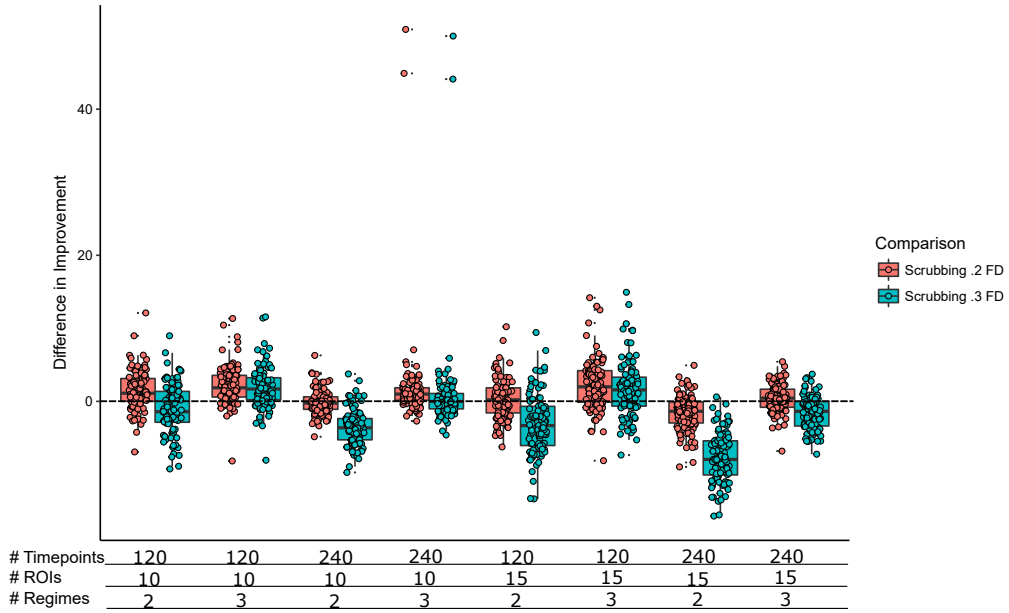


Figure 3.15: Comparison of the 2 regime CTRS-GARCH solution to Scrubbing at .2 and .3 FD within a simulation trial for Low Signal Discrete Motion Artifact Conditions. Negative numbers indicate that the CTRS-GARCH solution was superior to the scrubbing solution. These results indicate that the conservative scrubbing threshold of .2 was either equivalent or superior to 2 regime CTRS-GARCH in almost every condition, where as .3 FD scrubbing performed slightly worse than 2 regime CTRS-GARCH in almost every condition. As expected, the places were 2 regime CTRS-GARCH outperformed scrubbing were the high information conditions (240 time points, 2 regimes), however this improvement was small when compared to the high signal condition (3-5 units vs. 20-30 units.)

Comparing within a simulation iteration gives us a considerably clearer picture of how these motion correction techniques perform relative to one another. In the case of the high signal conditions, 2 regime CTRS-GARCH showed a significant improvement over .2 FD scrubbing in every condition (Paired Samples t-testing; $p < .05$ in all cases). While this significance level is in part due to the sample size (and therefore can be manipulated arbitrarily), it gives a certain amount of information. Overall, the majority of trials for every condition in the high signal condition set showed that 2 regime CTRS-GARCH gave a better estimate of the true correlation matrix than did .2 FD scrubbing. The worse performing conditions for 2 regime CTRS-GARCH were the 240 timepoint, 10 ROI and 3 regime condition, and the 120 time point, 15 ROI and 3 regime condition. While these conditions still showed a significant improvement over .2 FD scrubbing, there are still a considerable number of trials that showed scrubbing outperforming 2 regime CTRS-GARCH. Additionally, we see a large shift

when moving from .2 to .3 FD scrubbing, with 2 regime CTRS-GARCH considerably outperforming .3 FD scrubbing in all conditions.

The results for the Low signal condition set, in stark comparison to the High Signal condition set, suggest that .2 FD scrubbing either performs equally, or outperforms 2 regime CTRS-GARCH in almost every condition. Paired samples t-testing indicates that with the exception of the 240 time point, 10 ROI, 2 regime, the 120 time point, 15 ROI, 2 regime condition (both non-significant, $p = .41, .25$ respectively), and the 240 time point, 15 ROI, 2 regime condition (significant negative effect; $p < .05$), all other conditions showed a significant positive effect, suggesting that .2 FD scrubbing outperformed 2 regime CTRS-GARCH.

These results reaffirm the suggestions made based on the marginal results. If one has a large number of time points, and strong theoretical reason to believe that the underlying true signal will be strong, then 2 regime CTRS-GARCH will likely outperform scrubbing by far. However, if there are not many time points and if the signal is not thought to be very distinct, then the use of CTRS-GARCH runs the risk of leading to inaccurate estimates of the true correlation matrix, and at the very least does not perform as well as a conservative scrubbing approach.

3.2.2 Continuous Motion Artifact Results

Figures 3.16 and 3.17, again show the change in the absolute distance from the true signal matrix for the continuous motion artifact conditions. These are, again, marginal effects in the plots only show the improvement in estimation relative to the raw data estimate.

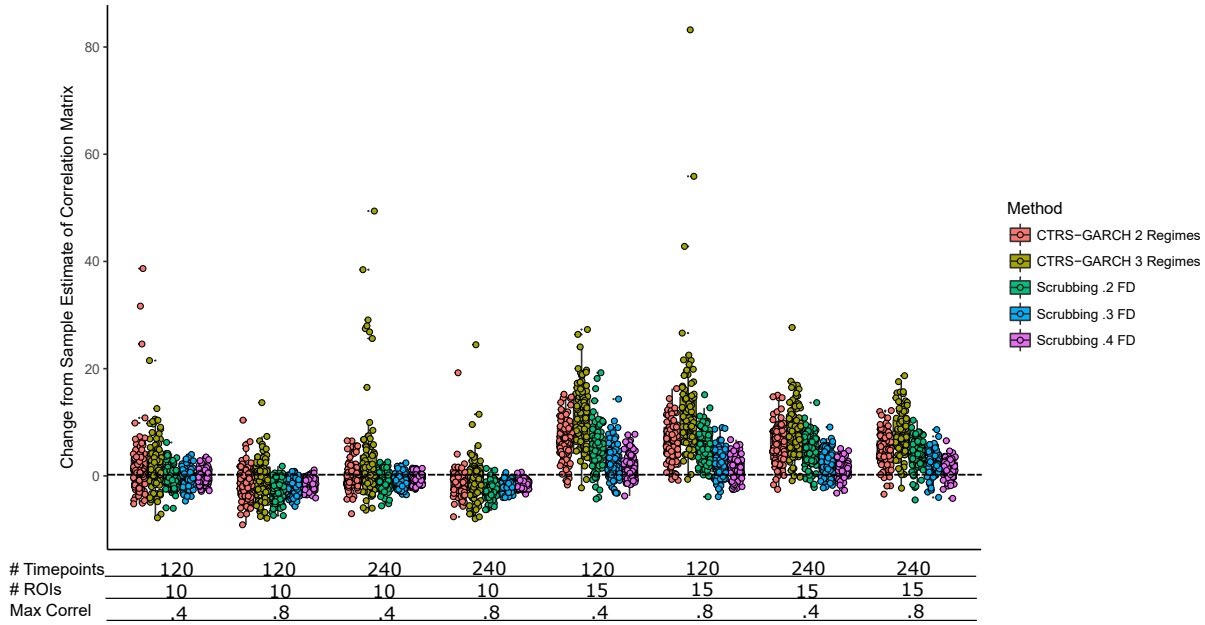


Figure 3.16: Change in the estimate of the true signal correlation matrix for the High Signal Continuous Motion Artifact Condition Set. Here, we see that in the 15 ROI condition set, all motion correction methods perform worse than the raw data estimate, while in the 10 ROI condition set, scrubbing is the only method that has the tendency to improve the estimation of the true correlation matrix. In all cases, the estimates from CTRS-GARCH are more variable than scrubbing, and are more likely to result in worse estimates of the true correlation matrix.

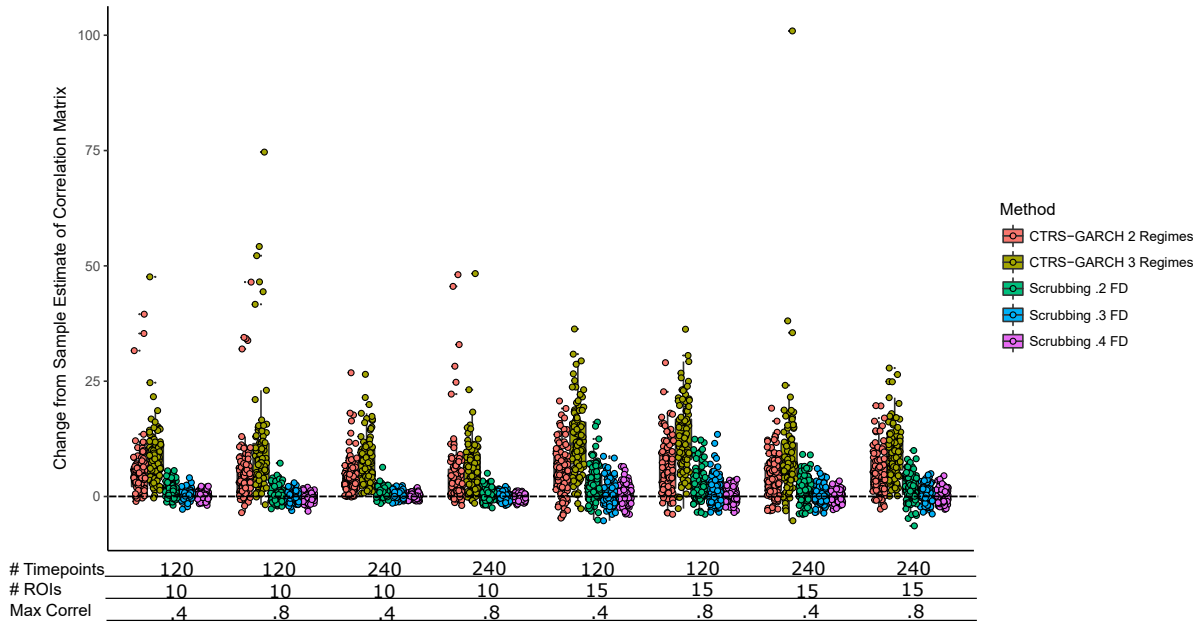


Figure 3.17: Change in the estimate of the true signal correlation matrix for the Low Signal Continuous Motion Artifact Condition Set. CTRS-GARCH, in all conditions, resulted in worse estimates of the true correlation matrix than the raw data estimates. Additionally, scrubbing appears to have no, or a slightly negative, effect on the quality of the estimate, as evidenced by the zero centering of the scrubbing distributions.

The results for the Continuous Motion Artifact Condition Set are markedly different from the discrete motion artifact conditions. Specifically, no motion correction technique evaluated here reliably improved the estimate of the true signal correlation matrix. There is one exception to this, and that is the High signal, 240 time points, 10 ROI and .8 maximum correlation condition. This condition is the most informative condition , and it shows that on average, all motion correction techniques improved the quality of the estimated true signal correlation matrix. However, this improvement was very slight (3-5 units).

Taken as a whole, these results suggest that a continuous motion artifact would be, in most cases, nearly impossible to correct with existing motion correction methods. CTRS-GARCH provides low quality estimates with high variability, and therefore would not be useful in correcting the motion artifact, while scrubbing tends to perform slightly better, but also has a tendency to result in lower quality estimates than simply using the raw data estimate.

In terms of general effects of design factors, we see the expected pattern of results. An increase in

the number of time points results in lower variance in the solutions, as well as a small improvement in quality, while having a stronger effect of motion results in a higher quality estimates from all motion correction methods. This is due to the effect of motion being more prominent, and therefore being easier to detect and extract. One unexpected result is the general effect of the number of ROIs. Increasing the number of ROIs from 10 to 15 results in a large decrease in the quality of the estimated true signal correlation matrix for all the motion correction methods.

Comparison of CTRS-GARCH and Scrubbing within iteration

Figures 3.18 and 3.19 show the relative performance of the 2 regime CTRS-GARCH solution compared to scrubbing at .2 FD and .3 FD.

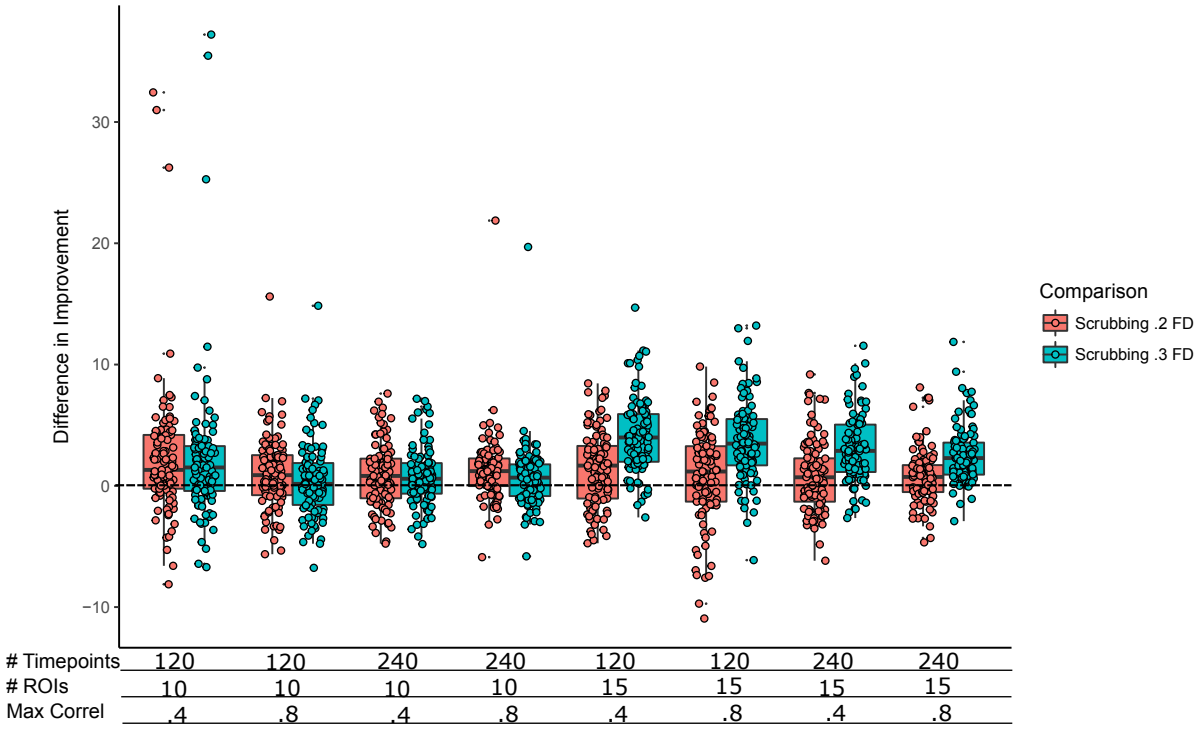


Figure 3.18: Comparison of the improvement over the uncorrected data from 2 Regime CTRS-GARCH to Scrubbing at .2 FD and .3 FD in the High Signal Continuous Motion Artifact Condition Set. Negative values indicate CTRS-GARCH performing better than the comparison. Note that in all cases, scrubbing out performs CTRS-GARCH. Notably, scrubbing at .3 FD tends to outperform CTRS-GARCH more than scrubbing at .2 FD. This is due to the previous findings that harsh scrubbing results in a lower quality estimate of the true signal correlation matrix than the uncorrected data estimate.

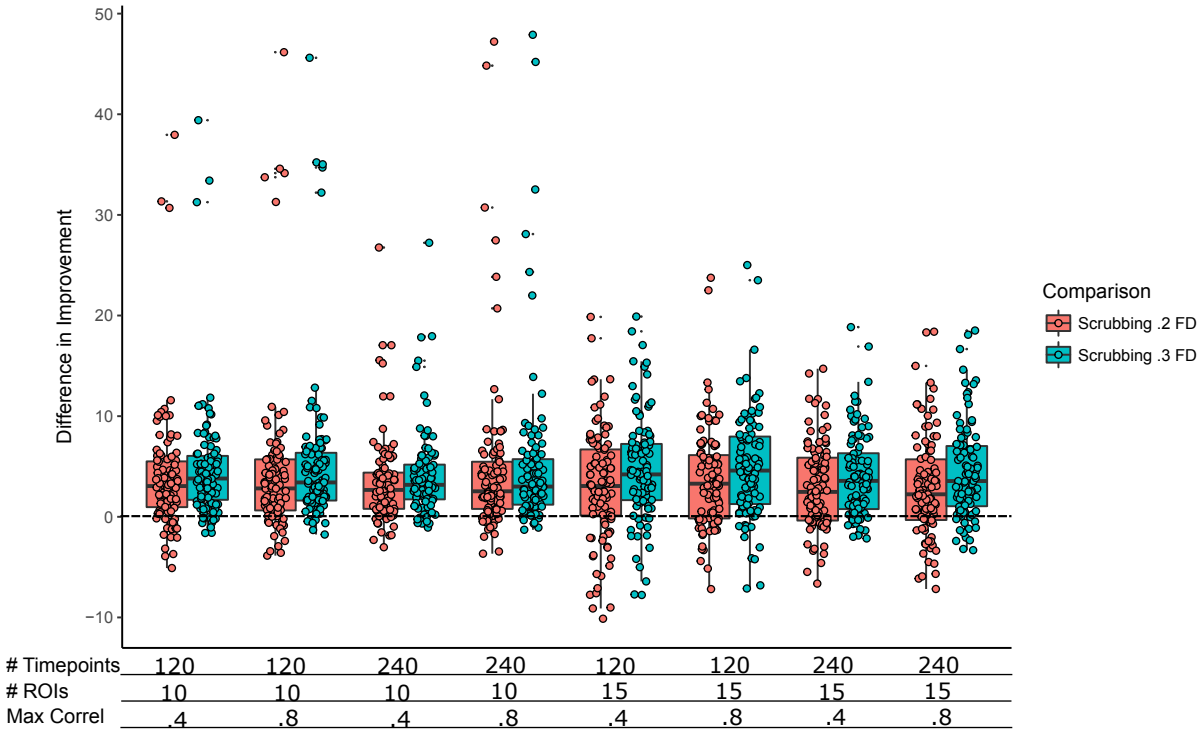


Figure 3.19: Comparison of the improvement over the uncorrected data from 2 Regime CTRS-GARCH to Scrubbing at .2 FD and .3 FD in the Low Signal Continuous Motion Artifact Condition Set. Negative values indicate CTRS-GARCH performing better than the comparison. Note a similar pattern of results as in the high signal condition, with the exception that scrubbing at .3 seems to be performing on par with scrubbing at .2.

The results for the paired comparisons of 2 regime CTRS-GARCH with scrubbing at .2 and .3 FD show a clear pattern for all conditions. In every case scrubbing performs better than CTRS-GARCH, and pairwise t-tests confirm this (in all cases $p < .05$). This result should be taken in context with the marginal results however. While scrubbing does perform better than CTRS-GARCH, it is the case that scrubbing tend to perform worse than estimates from uncorrected data. This is evidenced in the larger difference between scrubbing at .3 FD. This suggests that as the scrubbing threshold increases, and as the correct estimate approaches the uncorrected one, it approaches it from the incorrect direction. This is to say, scrubbing results in a worse estimate of the true signal correlation matrix than the uncorrected data.

3.2.3 Summary and Conclusion for Simulation Studies.

The results from the simulation studies show several unsurprising findings, and a great many surprising findings. To begin, the discrete motion artifact condition set showed for the most part the expected pattern of results. With higher information (more time points, fewer ROIs, larger shifts between regimes), CTRS-GARCH performed better than all other motion correction methods, and as information decreased, both the variance in performance as well as the mean level of performance of CTRS-GARCH decreased. This is to be expected as the discrete motion artifact condition is the data generating process implied by the CTRS-GARCH model. These sanity check simulations suggest that CTRS-GARCH is working as intended.

For the discrete motion artifact condition set, scrubbing also tended to improve the quality of the estimate of the true signal correlation matrix. Notably, scrubbing had the quality that it was rare any level of scrubbing caused a decrease in the quality of the estimate relative to the uncorrected data estimate. This suggests that scrubbing is safe to use regardless of the quality of the data at hand (albeit, this is only if motion artifacts follow a discrete data generation process).

The continuous motion artifact condition set tells a significantly different story. Here, no motion correction method had the tendency to remove the motion artifact, and in fact, all motion correction methods tended to make the quality of the estimate of the true signal correlation matrix worse. In part this is likely due to the constant nature of motion. If motion is always happening at a small level ($FD < .1$), a continuous motion artifact is always present, and there is no example of an uncontaminated set of data. As such, any estimate of a "true signal" correlation matrix will include some degree of motion contamination. However, it is telling that each motion correction method tended to increase the distance from the true correlation matrix, rather than simply not improve the estimate. This suggests that if the effect of motion is truly continuous in the fashion simulated here, the existing corrections for motion might be causing a large amount of bias in our functional connectivity estimates.

These simulation studies provide a simplified view of motion, and test CTRS-GARCHs performance in a controlled environment. Motion however, likely has a complex effect on functional connectivity and no simulation study no matter how complex will fully capture it. As such, we use an

empirical dataset in the next section to evaluate the performance of CTRS-GARCH and scrubbing in removing a real motion artifact.

4 ANALYSIS OF EMPIRICAL DATA

4.1 Overview

While the simulations from the previous section will allow us to evaluate the performance of the CTRS-GARCH algorithm writ large, the simulations rely on artificial data with very specific models of motion artifacts. There is no guarantee that the real data generating process behind the motion artifact is one of the simulated models, and rather it is more likely that the real data generating process is extremely complex. To that end, it is important to validate the CTRS-GARCH using a real fMRI dataset. Here, I evaluate the performance of the CTRS-GARCH, scrubbing and motion regression on a dataset of 166 participants resting state fMRI data. These 166 participants are children with mean age 11.41, and are divided into 4 groups of approximately 40, a typically developing control group, a group that has Autism Spectrum Disorder (ASD) diagnoses, a group with inattentive type ADHD diagnoses and a group with combined type ADHD diagnoses. More demographic and descriptive details of this data are contained in sections below. These data was generously provided by Dr. Damien Fair of Oregon Health and Science University, and received IRB approval from that institution prior to collection.

4.2 Analysis of Motion in Empirical Data

One of the key difficulties with using empirical data to evaluate the performance of a methodology is that one does not know the ground truth, which in this case is the true fMRI signal, un-distorted by motion. As such, we cannot rely on the methods that we used in the previous section. Here, we will use a different methodology, that of quality control-resting state functional connectivity correlations (QC-RSFC correlations), as proposed by Power et al. (2014). These correlations are simply the correlation between the mean Frame Displacements of the sample (calculated per individual) and each pairwise correlation between ROIs (again, calculated per individual). If we have, for example, 100 subjects evaluated with 5 ROIs, then we obtain $((5 \times 5) - 5)/2 = 10$ QC-RSDC correlations. We can evaluate QC-RSDC correlations before and after movement is extracted to determine if the

motion correction method is successful at removing the motion artifact. Ideally, after the motion correction method is applied, the average QC-RSFC correlation should be close to 0.

4.3 Data Overview

Demographics and descriptive statistics are contained in Table 4.1 and 4.2. In applied neuroimaging work at the level of ROI, it is typical to look at individual predefined networks. The networks used in this analysis will be the salience network, the dorsal attention network and the ventral attention network, with 6, 3 and 8 ROIs respectively, for 17 ROIs in total. Here, an automatic functional network parcellation based on the work of Power, Cohen, et al. (2011b) was used. These networks are chosen due to their association with autism (Menon & Uddin 2010; Uddin 2014) and ADHD (Castellanos & Proal 2012; Cortese et al. 2012). Each dataset contains 360 time points.

4.3.1 Data Acquisition and Preprocessing.

Individuals were scanned using a 3T Siemens Magnetom Tim Trio Scanner (Siemens, Erlangen, Germany) at OHSU Advanced Imaging Research Center. High resolution anatomical T1-weighted MPRAGE images were acquired in a sequence lasting 9 minutes and 14 seconds with the following parameters: TR = 2300 ms, TE = 3.58 ms, sagittal orientation, 256X256 voxels, 1^3 mm voxel dimensions. BOLD weighted functional data was collected using T2* weighted echo-planar imaging, with the following parameters: TR = 2500ms, TE = 30ms, flip angle = 90, FOV = 240mm, 36 slices interleaved, slice thickness 3.8 mm). Three resting state runs of 3.5 minutes were obtained.

Functional images were preprocessed with the following steps:

1. Removed MR signal offset spike.
2. Slice order correction.
3. Intensity normalization to mode value of 1000.
4. Functional data was coregistered to the anatomical MPRAGE, then coregistered to the Talarach atlas space.
5. The above coregistration corrects for gross head movements.

Finally, there are several fMRI signal preprocessing steps performed on the normalized images.

1. Band pass filter ($0.009Hz < f < 0.08Hz$)

2. Spatial smoothing (6 mm full width at half maximum)

3. Global signal regression

4. Ventricular and white matter regression.

No additional motion correction was applied.

Table 4.1: Frequency of Clinical Group by Gender. TDC is Typically Developing Control, ASD is Autism Spectrum Disorder, Inattentive is Inattentive type ADHD and Combined is Combined type ADHD

	TDC	ASD	Inattentive	Combined
Males	27	36	30	32
Females	15	7	11	9

Table 4.2: Demographic Descriptives

Demographic	Mean	SD
Age	11.41	1.75
IQ	106.43	15.46

4.4 Analysis and Results

To determine what effect each correction method had on the motion artifact, we used the QC-RSFC (quality control resting state functional connectivity) correlation method of Power et al. (2014). This method attempts to quantify the effect of motion in empirical functional connectivity data in the following fashion.

1. Begin by computing the correlation matrix using any motion correction method (or none) for each individual in your sample.

2. Extract the unique correlations (the upper or lower triangular) from each correlation matrix.

3. Arrange these in a dataset where rows are individuals, and columns are each unique correlation.

4. Compute mean Frame Displacement per Individual.

5. Correlate mean Frame Displacement with the Fisher Z transformed correlations.

6. These correlations of FD with functional connectivity are the QC-RSFC correlations.

QC-RSFC correlations have an intuitive interpretation. Theoretically, there should be no association of motion with the functional connectivity if motion has no effect on functional connectivity. Any correlation away from 0 indicates an increased level of motion contamination of that functional connectivity connection. This allows for motion to have different effects on different functional connectivity connections, as would be expected. Some FC relations might be uneffected by motion (as they might be located nearer to the center of motion, in the case of rotation), while others might be more effected by motion.

In this analysis the 17 ROIs used result in 136 QC-RSFC correlations. By calculating these per motion correction method, and then testing pairwise differences between the correlations (subject being the grouping factor), we can determine which motion correction method performs best empirically. Greater reduction in the QC-RSFC indicates that the motion correction method performs well.

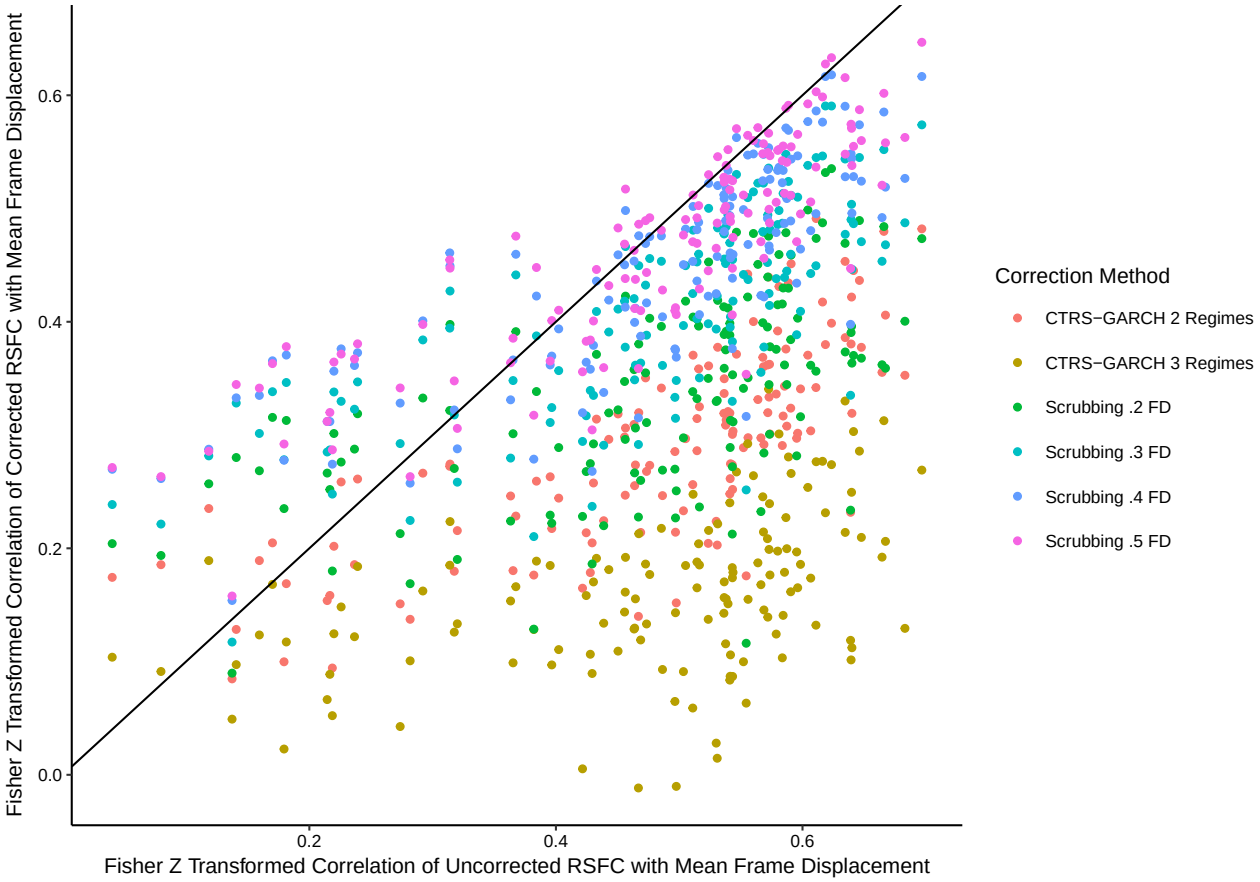


Figure 4.1: Plot of the Uncorrected QC-RSFCs with the Corrected QC-RSFCs for each motion correction method. The line represents no change in the QC-RSFCs. Below the line represents an improvement in QC-RSFCs while above represents a increase in the association with motion. Note that for higher uncorrected QC-RSFCs, scrubbing either improves or leaves unchanged the QC-RSFCs, while at low uncorrected QC-RSFCs, scrubbing increases the QC-RSFCs. Both CTRS-GARCH applications tend to improve QC-RSFCs dramatically.

Figure 4.1 above shows a scatter plot of the uncorrected QC-RSFCs with the corrected QC-RSFCs for each method under evaluation. Notably, scrubbing performs as expected at higher values of Uncorrected QC-RSFCs (greater impact of motion), in that it improves the QC-RSFCs in a matter inversely proportional to the severity of the threshold. In lower uncorrected QC-RSFCs scrubbing appears to exaggerate the effect of motion to a severe to moderate extent. This is likely due to lower uncorrected correlations being more related to noise. Figure 4.2 below shows only the plot for the two CTRS-GARCH corrections. Note that there appears to be a significant decrease in the QC-RSFC correlation with higher original QC-RSFC correlations for either the 2 regime or 3 regime CTRS-GARCH implementation, with 3 regime CTRS-GARCH performing considerably better than

2 regime CTRS-GARCH.

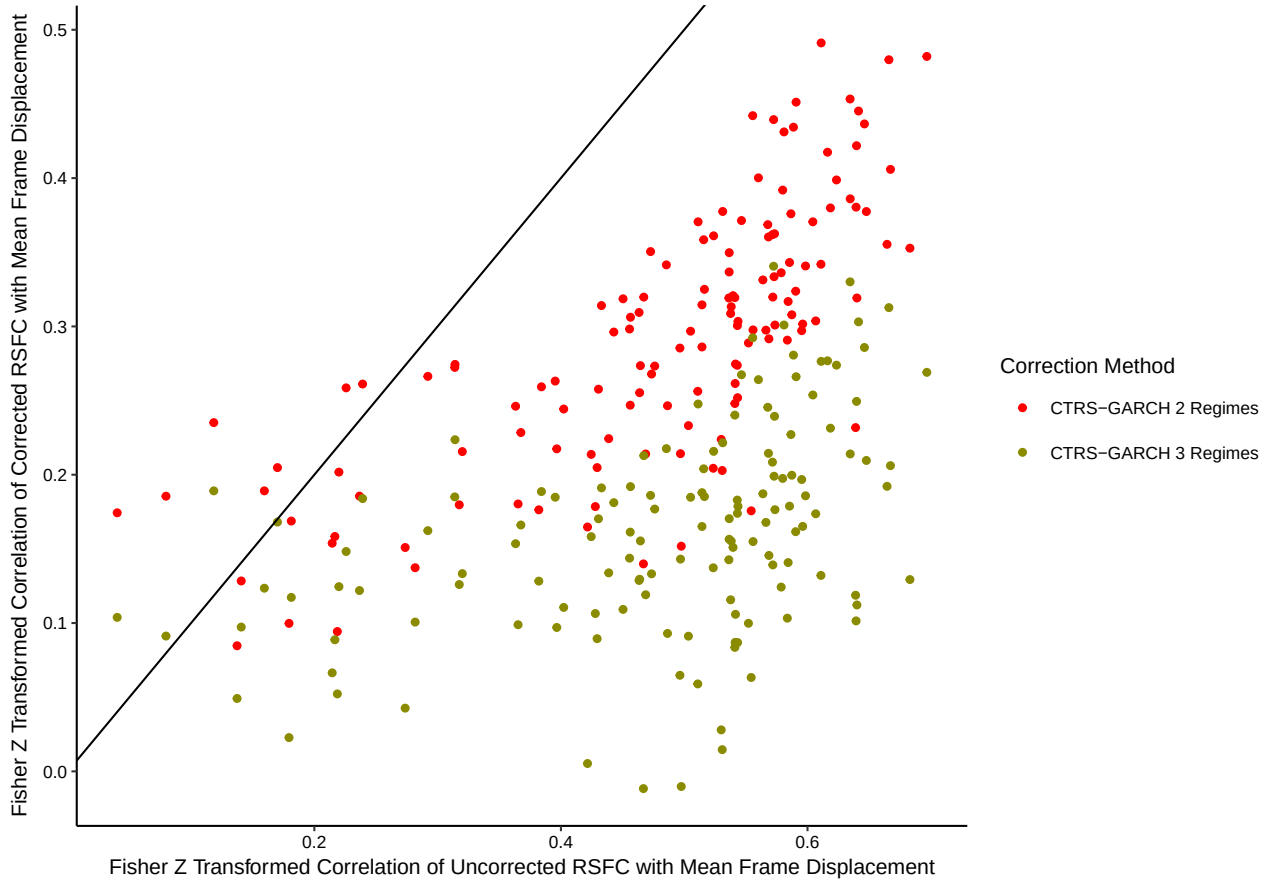


Figure 4.2: Plot of the Uncorrected QC-RSFCs with the Corrected QC-RSFCs for CTRS-GARCH methods. The line represents no change in the QC-RSFCs. Below the line represents an improvement in QC-RSFCs while above represents a increase in the association with motion.

CTRS-GARCH at 2 and 3 regimes tend to significantly outperform all versions of scrubbing. This difference is more prominent at higher levels of uncorrected QC-RSFC, but still present at lower levels. Notably, 3-regime CTRS-GARCH is the best performer, and even tends to reduce low uncorrected QC-RSFCs.

Table 4.3 shows the average within-subject differences in QC-RSFC correlations between the various motion correction methods, as well as the uncorrected data. Each of these differences were significant at an alpha level of .05 corrected using the Bonferroni correction.

Table 4.3: Overall pairwise mean differences between in the Fisher-Z transformed correlations of the RSFC-Correlations with mean frame displacement. All differences were significant at the .05 level with a Bonferroni correction for multiple comparisons. Negative values mean that the row method reduced the QC-RSFC correlation more than the column method. Note that 3 Regime CTRS-GARCH outperforms all other methods, including 2 Regime CTRS-GARCH.

	2 Regimes	3 Regimes	FD .2	FD .3	FD .4	FD .5
3 Regimes	-0.119					
FD .2	0.03	0.15				
FD .3	0.113	0.232	0.082			
FD .4	0.155	0.274	0.124	0.041		
FD .5	0.178	0.298	0.148	0.065	0.023	
Uncorrected	0.206	0.326	0.176	0.093	0.051	0.027

There are three salient details regarding these results. The first is that on average, all motion correction methods tended to decrease the QC-RSFC, and presumably remove a portion of the effect of motion. Scrubbing showed the expected pattern of results, where lower thresholds for scrubbing resulted in more removal of the motion artifact, while higher thresholds resulted in QC-RSFC closer to the raw data correlation. The second point is that CTRS-GARCH with 2 regimes performs slightly better than the most severe form of scrubbing (.2 FD). This difference is slight (.03) but highly significant ($p < .001$). The final point is that 3 regime CTRS-GARCH outperforms all other motion correction methods by a fairly large amount, ranging from .15 improvement over .2 FD scrubbing, and .3 at .5 FD scrubbing.

4.4.1 Between Diagnostic Category Differences.

Here, we examine any differences between diagnostic categories in terms of their responses to the motion correction techniques. To evaluate this, we recalculate the QC-RSFC correlations within each diagnostic category, and apply a 2 way repeated measures ANOVA, evaluating differences in the within-correlation factor of motion correction, and the additional within correlation factor of diagnostic category.

The results from the 2-way repeated measures ANOVA show significant effects for both the motion correction method and the diagnostic category ($F(6, 810) = 759.80, p < .05$ and $F(3, 408) = 134.61, p < .05$) as well as a significant interaction ($F(18, 2430) = 25.28, p < .05$). To parse apart these findings, we can inspect a plot of marginal effects. Figure 4.3 below shows a plot of the average change from uncorrected QC-RSFC for each diagnostic group and each motion correction method.

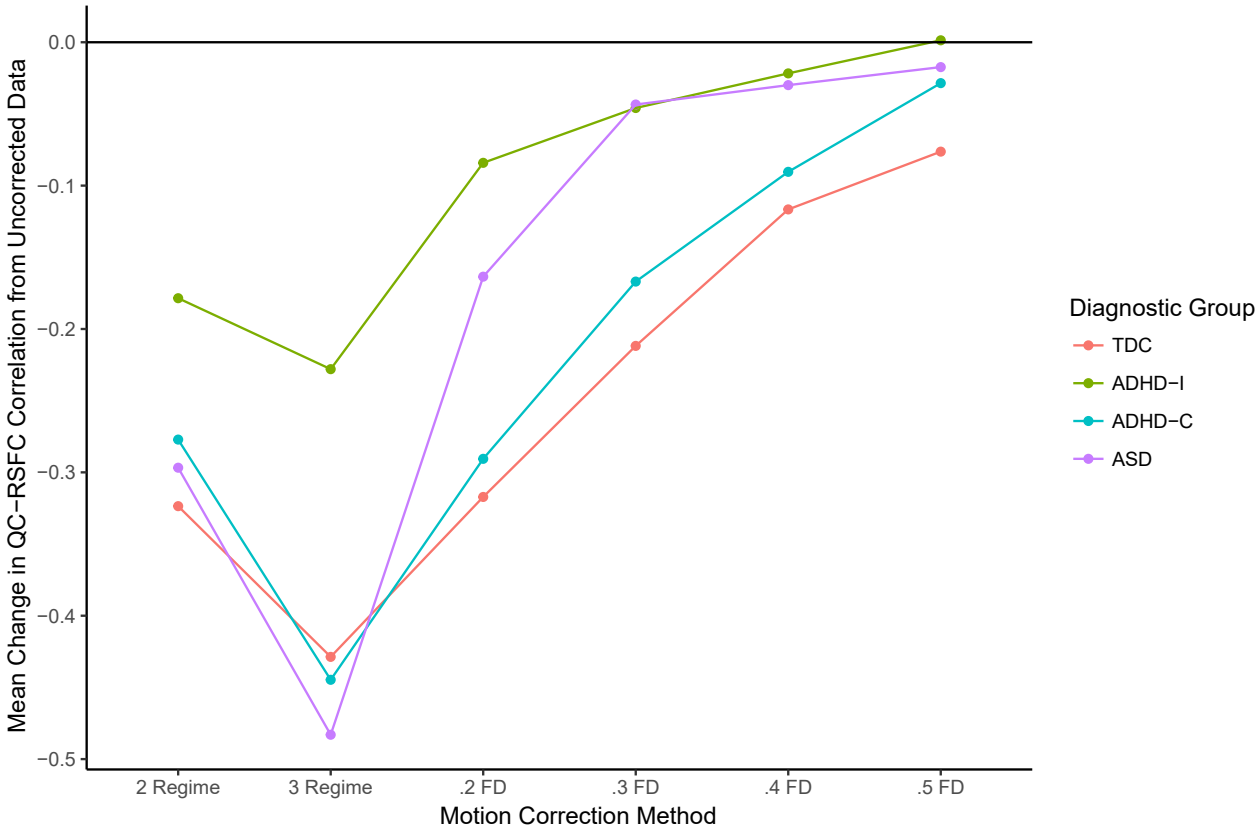


Figure 4.3: The average change from the uncorrected QC-RSFC (Fisher Z transformed) by diagnostic category and motion correction method (.2 FD is scrubbing at .2 FD, 2 regimes is 2 regime CTRS-GARCH, etc.). Note the similarities in effect for the TDC, ASD and ADHD-C, and that the main differences are in the ADHD-I group.

Each diagnostic category shows a similar pattern of results. In each case, the best improvement is found in the 3 regime CTRS-GARCH, and improvement lessens as the scrubbing threshold increases. The main finding of Figure 4.3 is that the ADHD-I subgroup shows a significantly lower improvement than any other subgroup for the majority of the motion correction methods. This is a curious finding, and not one due to lower motion in the ADHD-I group. A 1-way ANOVA was used to assess differences in the mean frame displacement between diagnostic categories, and was found

to be non-significant ($F(3, 154) = 2.002, p = .116$). Figures 4.4 and 4.5 below show the uncorrected QC-RSFC values compared to corrected QC-RSFC values for only the ADHD-I subgroup, and for only the TDC subgroup respectively. The axes are equal, allowing for visual comparison.

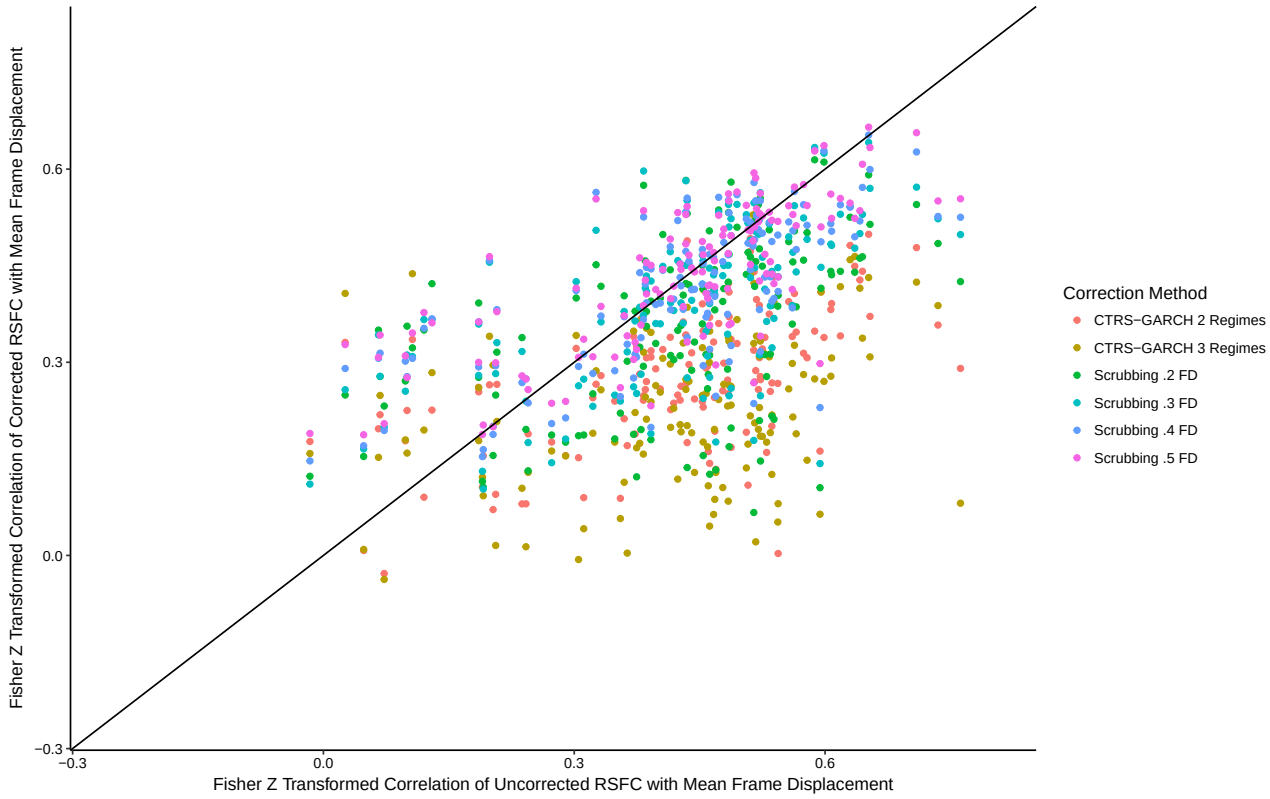


Figure 4.4: QC-RSFC comparisons for the ADHD-I subgroup. Note the larger number of points above the diagonal line, denoting an increase in the correlation with motion.

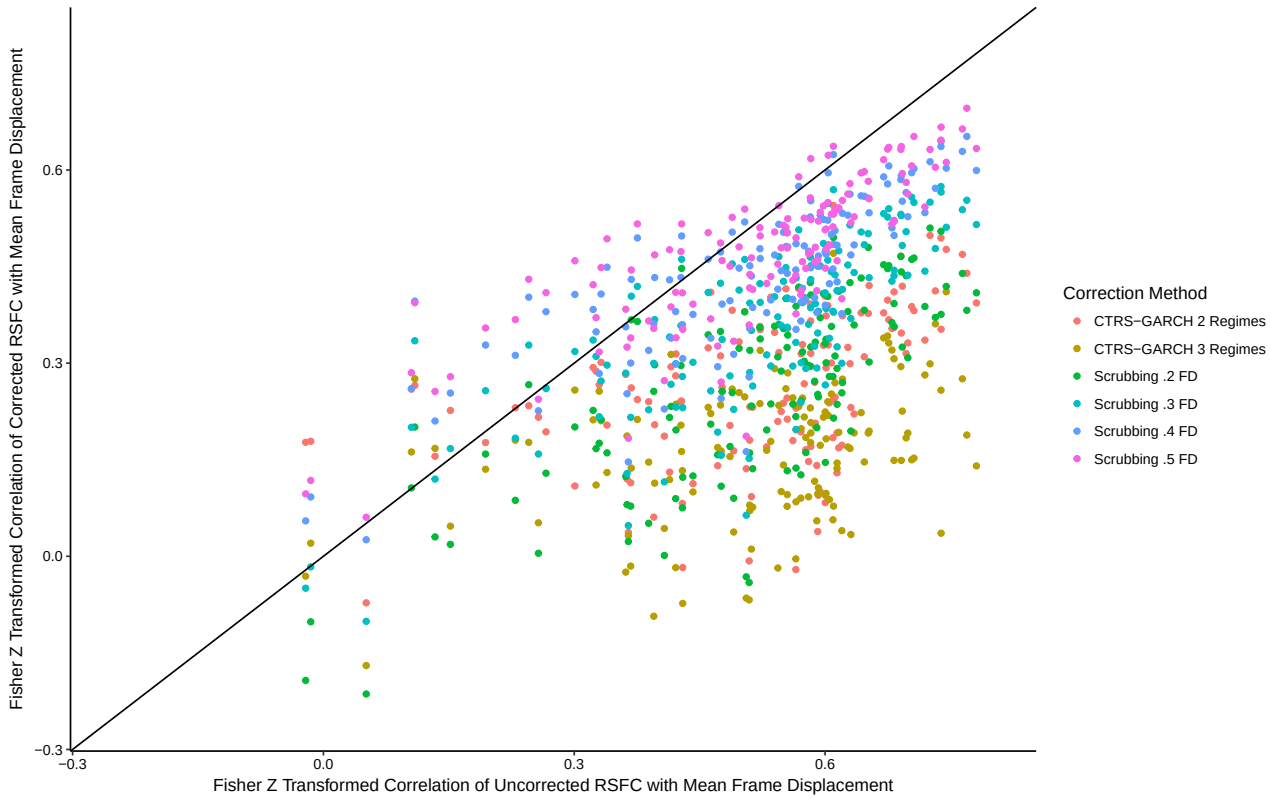


Figure 4.5: QC-RSFC comparisons for the TDC subgroup. Note the larger range of correlations in comparison with the ADHD-I, as well as a greater concentration of those correlations in the upper area of the uncorrected QC-RSFC.

Inspection of the figures above suggests that the overall difference in performance of motion correction methods within the ADHD-I subgroup is due to the clustering of more uncorrected QC-RSFC correlations in the mid to low range. From the overall group plots, we know that the higher the uncorrected QC-RSFC correlation is, the more a motion correction method is going to correct it. This suggests that the ADHD-I subgroup was less effected by motion than the other groups, while still having on average the same amount of motion. While this is an interesting finding, a substantive explanation of the phenomena is beyond the scope of this dissertation. That being said, even with the decrease in performance for that single subgroup, CTRS-GARCH with 3 regimes outperforms all other motion correction methods when applied to this dataset.

4.5 Summary

The empirical results show that CTRS-GARCH tends to outperform scrubbing at removing the association of functional connectivity with motion, often times to a large degree. This effect is most

prominent using a 3 regime solution, but is also present in a 2 regime solution. These differences are statistically significant in all cases. Analysis of each diagnostic category showed that for the ASD, ADHD-C and TDC subgroups, motion correction methods tended to have the same effect. For the ADHD-I subgroup, the improvement in the QC-RSFC correlations was less than for other subgroups. This was likely due to there being lower uncorrected QC-RSFC for the ADHD-I subgroup, and therefore a reduced need for correction. However, this implies differences in how motion effects this subgroup, a substantive question that this dissertation is unable to answer.

5 DISCUSSION

This dissertation developed the conditional transition regime switching generalized autoregressive conditional heteroskedasticity model, or CTRS-GARCH for short. This model was developed specifically for the purposes of correcting motion during functional MRI scans of high value subjects, such as pediatric subjects. CTRS-GARCH was developed as an improvement on the existing best performing motion correction method, scrubbing or the removal of motion contaminated time points, in that CTRS-GARCH is not necessarily destructive, as it does not remove time points wholesale. What CTRS-GARCH attempts to do is automatically identify motion contaminated time points, remove the effect of motion from that time point, and return an estimate of a motion free correlation matrix. It achieves this by using regime-switching GARCH to model unobserved changes in the correlation structure, and ties these changes to the observed motion of the subject through a conditional transition component.

This dissertation develops the estimation framework, integrates the novel contribution of conditional transitions with what is effectively an ordered logistic modeling strategy, and evaluates the performance of CTRS-GARCH against other motion correction strategies using two sets of simulations and an empirical dataset of 160 resting state scans of children with ASD, ADHD-C, ADHD-I and typically developing controls.

The simulations presented here consisted of a discrete motion artifact generation condition, and a continuous motion artifact generation condition. Within the discrete motion artifact condition, the results indicated that for functional connectivity networks with high signal, CTRS-GARCH retrieves the true signal correlation matrix better than scrubbing. In simulations with fewer time points and/or more ROIs, the performance of the CTRS-GARCH algorithm decreases, but tends to perform on par with scrubbing. One notable finding from the discrete motion artifact generation condition was that 2 regime CTRS-GARCH, wherein there is a single regime for the true correlation matrix, and one regime for the effect of motion, was the best performer, while 3 regime CTRS-GARCH (2 motion

regimes) had significantly higher variance in its' solutions and had the tendency to result in worse estimates of the true signal correlation matrix than an uncorrected, or raw, data estimate.

The continuous motion artifact generation condition was considerably less kind on all the motion correction methods than then discrete motion correction condition. In the large majority of cases, any motion correction technique tended to result in worse estimates of the true signal correlation matrix than an uncorrected estimate. The one exception was in the highest information condition (High signal, highest number of time points, lowest number of ROIs), where scrubbing tended to improve the estimate of the true signal correlation matrix by a minor amount. In all cases, scrubbing was less detrimental to the estimate of the true signal correlation than CTRS-GARCH was.

The results of the simulation studies show two very different scenarios. In the first case, motion can be corrected for and CTRS-GARCH tends to perform well. In the second case, no motion correction technique does well. This state of affairs would be difficult to conclude on, but fortunately, this dissertation also contains an empirical analysis which suggests that CTRS-GARCH performs quite well at removing the motion artifact in a real-world setting.

In the empirical analysis, we used resting state data from approximately 160 children evenly divided into children with autism spectrum disorder, attention deficit/hyperactivity disorder combined, attention deficit/hyperactivity disorder inattentive, and typically developing controls. Using the methodology of (Power et al. 2014), we analyzed the association of motion with functional connectivity results using the quality control - resting state functional connectivity correlations. These correlations are an indicator of how associated is a particular functional relation to average motion. A non-zero QC-RSFC indicates motion contamination. We applied CTRS-GARCH with 2 and 3 regimes, as well as scrubbing at .2, .3, .4, and .5 to the datasets in this study and analyzed the change in the QC-RSFC. These results show that 3-regime CTRS-GARCH resulted in the most reduction in the QC-RSFC out of any of the motion correction methods. One interesting finding was that at low uncorrected QC-RSFC motion correction methods tended to increase the QC-RSFC. This is likely due to low QC-RSFC being a result of noise rather than a true relation with motion. Finally, an analysis of the diagnostic subgroups show that the ADHD-I subgroup had significantly less reduction in their QC-RSFCs than any other subgroup, even though they were no different in terms of their mean

frame displacement. This is an interesting finding, but one that this dissertation has no explanation for.

There are several limitations to this study. The first is that due to the complex nature of the motion artifact, the simulations presented here cannot be thought of as fully representative. Indeed, given the opposite findings between the discrete motion artifact and the continuous motion artifact, these results can only provide a guide and a word of caution as to the use of CTRS-GARCH (and scrubbing.) This is in part alleviated by the empirical example, which shows CTRS-GARCH outperforming scrubbing by a considerable margin. Yet, there is still a question of what the motion artifact form is, and if QC-RSFC, and the associated change in QC-RSFC, is related to a reduction in the motion artifact, or if we are moving even farther away from the truth. The QC-RSFC implies a linear relation between absolute movement (in the form of frame displacement) and the value of the correlation between two ROIs. This is quite a restrictive model, and doesn't take into account the implied dependence structure of two correlations in the same matrix. This could be a fruitful further avenue of research, determining high quality metrics of motion contamination in fMRI data, though this author has no suggestions at this time as to obtaining these.

Overall these results suggest two things. The first is that CTRS-GARCH has the potential to be a highly effective motion correction method. The second is that scrubbing continues to be a highly reliable method of motion correction. The mechanics and effects of scrubbing are well known, and have proven to be effective at reducing the motion artifact in countless studies. However, as was one of the central points of this dissertation, scrubbing is quite conservative. CTRS-GARCH can be thought of as a parameterized soft thresholding (in that not all of a data point is removed), with a subject specific thresholding "point". Indeed, CTRS-GARCH was designed to extend and complement a scrubbing approach. The results of this dissertation, both simulation as well as empirical, suggest that substantive researchers apply scrubbing at several thresholds as well as use CTRS-GARCH at several numbers of regimes as a way of performing a sensitivity analysis with their study.

There are several future directions to go with regards to this approach. The first is refining the data generation procedure for fMRI so that it better matches the actual physical process. This would

involve simulating what would effectively be an artificial subject and scanner, simulate the data acquisition procedure, and from physical principles, simulate the effect of motion on that data acquisition. This would allow researchers to have a better test bed for developing methods to analyze fMRI data. Another direction is to refine the regime switching GARCH to simultaneously extract motion, as well as model dynamic functional connectivity. This would allow the results from CTRS-GARCH to more accurately reflect the dynamic nature of functional connectivity. Finally, an experimental investigation of the effect of motion is in order. By systematically manipulating subjects in a scanner, one would be able to determine the effect of motion on functional connectivity. This would involve isolating head movements of the subject, and manipulating the subject's head to move in a set fashion. This would increase our understanding of the motion artifact considerably.

In closing, motion continues to be a pernicious foe to the fMRI researcher. The work here is presented in the hopes that it can help the substantive researcher extract more value out of high value, high motion samples that often form the datasets behind highly salient clinical problems.

REFERENCES

Berndt, E. K., Hall, B. H., Hall, R. E., & Hausman, J. (1974). Estimation and Inference in Nonlinear Structural Models. *October*, *3*(4), 653 - 665.

Biswal, B., Yetkin, F. Z., Haughton, V. M., & Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echoplanar mri. *Magnetic resonance Medicine*(9), 537–541.

Bollen, K. A. (1989). *Structural Equation Models with Latent Variables* (Vol. 9) (No. 5). Wiley. doi: 10.1177/1461444807080339

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*(3), 307–327. doi: 10.1016/0304-4076(86)90063-1

Bollerslev, T., Engle, R., & Wooldridge, M. (1988). A Capital Asset Pricing Model with Covariances. *Journal of Political Economy*, *96*(11), 116–131.

Brodmann, K. (1909). *Vergleichende Lokalisationslehre der Großhirnrinde: in ihren Prinzipien dargestellt auf Grund des Zellenbaues*. Leipzig: Johann Ambrosius Barth. doi: 10.1097/00005053 -191012000-00013

Bullmore, E. T., & Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature reviews. Neuroscience*, *10*(3), 186–98. doi: 10.1038/nrn2575

Castellanos, F. X., & Proal, E. (2012). Large-scale brain systems in ADHD: beyond the prefrontals-triatal model. *Trends in Cognitive Sciences*, *16*(1), 17–26. doi: 10.1016/j.tics.2011.11.007

Cortese, S., Kelly, C., Chabernaud, C., Proal, E., Di Martino, A., Milham, M. P., & Castellanos, F. X. (2012). Toward Systems Neuroscience of ADHD: A Meta-Analysis of 55 fMRI Studies. *American Journal of Psychiatry*, *169*(10), 1038–1055. doi: 10.1176/appi.ajp.2012.11101521

Craddock, R. C., James, G. A., Holtzheimer, P. E., Hu, X. P., & Mayberg, H. S. (2012). A whole brain fMRI atlas generated via spatially constrained spectral clustering. *Human Brain Mapping*, *33*(8), 1914–1928. doi: 10.1002/hbm.21333

Dempster, a. P., Laird, N. M., & Rubin, D. B. (1977). Maximum Likelihood from incomplete Data via the EM Algorithm. *Journal of the Royal Statistica Society*, *39*(1), 1–38. doi: 10.1.1.133.4884

Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., ... Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, *31*(3), 968–980. doi: 10.1016/j.neuroimage.2006 .01.021

Diebold, F. X., Lee, J.-H., & Weinbach, G. C. (1994). Regime switching with time-varying transition probabilities. *Nonstationary Time Series Analysis and Cointegration*, 283–302.

Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of

United Kingdom Inflation. *Econometrica*, *50*(4), 987–1007. doi: 10.2307/1912773

Engle, R. F. (2001). GARCH 101: The Use of ARCH/GARCH Models in Applied Econometrics. *Journal of Economic Perspectives*, *15*(4), 157–168. doi: 10.1257/jep.15.4.157

Engle, R. F. (2002). Dynamic Conditional Correlation. *Journal of Business & Economic Statistics*, *20*(3), 339–350. doi: 10.1198/073500102288618487

Erhardt, E. B., Allen, E. A., Wei, Y., Eichele, T., & Calhoun, V. D. (2012). SimTB, a simulation toolbox for fMRI data under a model of spatiotemporal separability. *NeuroImage*, *59*(4), 4160–4167. doi: 10.1016/j.neuroimage.2011.11.088

Fair, D. A., Nigg, J. T., Iyer, S., Bathula, D., Mills, K. L., Dosenbach, N. U. F., ... Milham, M. P. (2012). Distinct neural signatures detected for ADHD subtypes after controlling for micro-movements in resting state functional connectivity MRI data. *Frontiers in systems neuroscience*, *6*(February), 80. doi: 10.3389/fnsys.2012.00080

Friston, K. J., Harrison, L., & Penny, W. (2003). Dynamic causal modelling. *NeuroImage*, *19*(4), 1273–1302. doi: 10.1016/S1053-8119(03)00202-7

Friston, K. J., Williams, S., Howard, R., Frackowiak, R. S., & Turner, R. (1996). Movement-related effects in fMRI time-series. *Magnetic resonance in medicine : official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, *35*(3), 346–355. doi: DOI10.1002/mrm.1910350312

Gomot, M., Bernard, F., Davis, M., Belmonte, M., Ashwin, C., Bullmore, E., & Baron-Cohen, S. (2006). Change detection in children with autism: an auditory event-related fMRI study. *NeuroImage*, *29*(2), 475.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series. *Econometrica*, *57*(2), 357–384. doi: 10.2307/1912559

Hamilton, J. D. (1990). Analysis of time series subject to changes in regime. *Journal of Econometrics*, *45*(1-2), 39–70. doi: 10.1016/0304-4076(90)90093-9

Hamilton, J. D. (1994). *Time Series Analysis*. Princeton, New Jersey: Princeton University Press.

Hindriks, R., Adhikari, M. H., Murayama, Y., Ganzetti, M., Mantini, D., Logothetis, N. K., & Deco, G. (2016). Can sliding-window correlations reveal dynamic functional connectivity in resting-state fMRI? *NeuroImage*, *127*, 242–256. doi: 10.1016/j.neuroimage.2015.11.055

Huettel, S., Song, A., & McCarthy, G. (2014). *Functional Magnetic Resonance Imaging* (3rd ed.). Sinauer Associates, Inc.

Hutchison, R. M., Womelsdorf, T., Allen, E. A., Bandettini, P. A., Calhoun, V. D., Corbetta, M., ... Chang, C. (2013). Dynamic functional connectivity: Promise, issues, and interpretations. *NeuroImage*, *80*, 360–378. doi: 10.1016/j.neuroimage.2013.05.079

Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, *17*(2), 825–841. doi: 10.1016/S1053-8119(02)91132-8

Kim, C.-J., & Nelson, C. R. (1999). *State-space models with regime switching : classical and Gibbs-sampling approaches with applications.* Boston: The MIT Press. doi: 10.2307/2669796

Kim, J., Zhu, W., Chang, L., Bentler, P. M., & Ernst, T. (2007). Unified structural equation modeling approach for the analysis of multisubject, multivariate functional MRI data. *Human Brain Mapping*, *28*(2), 85–93.

Lindquist, M. A., Xu, Y., Nebel, M. B., & Caffo, B. S. (2014). Evaluating dynamic bivariate correlations in resting-state fMRI: A comparison study and a new approach. *NeuroImage*, *101*(4), 531–546. doi: 10.1016/j.neuroimage.2014.06.052

Logothetis, N. (2008). What we can, and what we can't do with fMRI. *Nature*, *453*(7197), 869–879.

Ma, S., Calhoun, V. D., Phlypo, R., & Adali, T. (2014). Dynamic changes of spatial functional network connectivity in healthy individuals and schizophrenia patients using independent vector analysis. *NeuroImage*, *90*, 196–206. doi: 10.1016/j.neuroimage.2013.12.063

Menon, V., & Uddin, L. Q. (2010). Saliency, switching, attention and control: a network model of insula function. *Brain Structure and Function*, 1–13. doi: 10.1007/s00429-010-0262-0

Molenaar, P. C., Beltz, A. M., Gates, K. M., & Wilson, S. J. (2016). State space modeling of time-varying contemporaneous and lagged relations in connectivity maps. *NeuroImage*, *125*, 791–802. doi: 10.1016/j.neuroimage.2015.10.088

Murphy, K., Birn, R. M., Handwerker, D. A., Jones, T. B., & Bandettini, P. A. (2009). The impact of global signal regression on resting state correlations: Are anti-correlated networks introduced? *NeuroImage*, *44*(3), 893–905. doi: 10.1016/j.neuroimage.2008.09.036

Ogawa, S., Lee, T. M., Kay, A. R., & Tank, D. W. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences of the United States of America*, *87*(24), 9868–72. doi: 10.1073/pnas.87.24.9868

Pelletier, D. (2006). Regime switching for dynamic correlations. *Journal of Econometrics*, *131*(1-2), 445–473. doi: 10.1016/j.jeconom.2005.01.013

Penny, W., Friston, K., Ashburner, J., & Nichols, T. (2006). *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. London: Elsevier.

Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2011). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage*, *59*(3), 2142–2154. doi: 10.1016/j.neuroimage.2011.10.018

Power, J. D., Cohen, A. L., Nelson, S. M., Wig, G. S., Barnes, K. A., Church, J. A., ... Petersen, S. E. (2011a). Functional Network Organization of the Human Brain. *Neuron*, *72*(4), 665–678.

doi: 10.1016/j.neuron.2011.09.006

Power, J. D., Cohen, A. L. L., Nelson, S. M. M., Wig, G. S. S., Barnes, K. A. A., Church, J. A. A., ... Petersen, S. E. E. (2011b). Functional network organization of the human brain. *Neuron*, *72*(4), 665–78. doi: 10.1016/j.neuron.2011.09.006

Power, J. D., Mitra, A., Laumann, T. O., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2014). Methods to detect, characterize, and remove motion artifact in resting state fMRI. *NeuroImage*, *84*, 320–341. doi: 10.1016/j.neuroimage.2013.08.048

Saad, Z. S., Gotts, S. J., Murphy, K., Chen, G., Jo, H. J., Martin, A., & Cox, R. W. (2012). Trouble at rest: how correlation patterns and group differences become distorted after global signal regression. *Brain connectivity*, *2*(1), 25–32. doi: 10.1089/brain.2012.0080

Samejima, F. (1969). Graded Response Model. In *Handbook of modern item response theory.*

Satterthwaite, T. D., Elliott, M. A., Gerraty, R. T., Ruparel, K., Loughead, J., Calkins, M. E., ... Wolf, D. H. (2013a). An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. *NeuroImage*, *64*(1), 240–256. doi: 10.1016/j.neuroimage.2012.08.052

Satterthwaite, T. D., Elliott, M. A., Gerraty, R. T., Ruparel, K., Loughead, J., Calkins, M. E., ... Wolf, D. H. (2013b). An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. *NeuroImage*, *64*(1), 240–256. doi: 10.1016/j.neuroimage.2012.08.052

Satterthwaite, T. D., Wolf, D. H., Loughead, J., Ruparel, K., Elliott, M. A., Hakonarson, H., ... Gur, R. E. (2012). Impact of in-scanner head motion on multiple measures of functional connectivity: Relevance for studies of neurodevelopment in youth. *NeuroImage*, *60*(1), 623–632. doi: 10.1016/j.neuroimage.2011.12.063

Smith, E. B., Menon, T., & Thompson, L. (2011). Status Differences in the Cognitive Activation of Social Networks. *Organization Science*, *23*(1), 67–82. doi: 10.1287/orsc.1100.0643

Smyser, C. D., Inder, T. E., Shimony, J. S., Hill, J. E., Degnan, A. J., Snyder, A. Z., & Neil, J. J. (2010). Longitudinal analysis of neural network development in preterm infants. *Cereb Cortex*, *20*(12), 2852–2862. doi: 10.1093/cercor/bhq035

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., ... Joliot, M. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage*, *15*(1), 273–289. doi: 10.1006/nimg.2001.0978

Uddin, L. Q. (2014). Salience processing and insular cortical function and dysfunction. *Nature Reviews Neuroscience*, *16*(1), 55–61. doi: 10.1038/nrn3857

van Dijk, K. R. A., Sabuncu, M. R., & Buckner, R. L. (2011). The influence of head motion on intrinsic functional connectivity MRI. *NeuroImage*, *59*(1), 431–438. doi: 10.1016/j.neuroimage

.2011.07.044