2-1-2016

Development of a Group Dynamic Functional Connectivity Pipeline for Magnetoencephalography Data and its Application to the Human Face Processing Network

Per Lysne

Recommended Citation

Lysne, Per. "Development of a Group Dynamic Functional Connectivity Pipeline for Magnetoencephalography Data and its Application to the Human Face Processing Network." (2016). https://digitalrepository.unm.edu/psy_etds/85

Per A. Lysne

Candidate

Psychology

Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

Harold D. Delaney, Chairperson

Claudia D. Tesche, co-Chairperson

Kristina T. Ciesielski

Robert J. Thoma

DEVELOPMENT OF A GROUP DYNAMIC FUNCTIONAL CONNECTIVITY PIPELINE FOR MAGNETOENCEPHALOGRAPY DATA AND ITS APPLICATION TO THE HUMAN FACE PROCESSING NETWORK

by

PER A. LYSNE

B.S. Computer Engineering
M.S. Electrical Engineering
M.S. Psychology

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
Psychology

The University of New Mexico
Albuquerque, New Mexico

December, 2015

Dedication

This dissertation is dedicated to the animals who have lost their lives in research.

Acknowledgements

I am pleased to generously acknowledge the members of my dissertation committee, Drs. Harold Delaney, Claudia Tesche, Kristina Ciesielski, and Robert Thoma. Thank you for your patience with the time it took to complete this research, and then with my final sprint to the finish. Particular thanks to Robert Thoma for his friendship and support while I learned neuroimaging, and to Harold Delaney for his mentorship and guidance in completing graduate school. Bob, you were right, it is harder than it looks. Harold, your example is one I aspire to emulate.

Thank you to my family, Peter, Jeanette, and JoAnn Lysne, Patrick McCormick, and Elizabeth Jaenicke, for your encouragement and patience with this process. Thank you to my grandparents, Osborne and Louise Lysne, and Arnold and Hazel Hansen, for your love and encouragement to pursue education. You are not forgotten.

Development of a Group Dynamic Directed Functional Connectivity Pipeline for Magnetoencephalography Data and its Application to the Human Face Processing Network

By

Per A. Lysne

B.S. Computer Engineering, University of New Mexico, 1990
M.S. Electrical Engineering, University of Arizona, 1996
M.S. Psychology, University of New Mexico, 2009
Ph.D. Psychology, University of New Mexico, 2015

Abstract

Since its inception, functional neuroimaging has focused on identifying sources of neural activity. Recently, interest has turned to the analysis of connectivity between neural sources in dynamic brain networks. This new interest calls for the development of appropriate investigative techniques.

A problem occurs in connectivity studies when the differing networks of individually analyzed subjects must be reconciled. One solution, the estimation of group models, has become common in fMRI, but is largely untried with electromagnetic data. Additionally, the assumption of stationarity has crept into the field, precluding the analysis of dynamic systems. Group extensions are applied to the sparse irMxNE localizer of MNE-Python. Spectral estimation requires individual source trials, and a multivariate multiple regression procedure is established to accomplish this based on the irMxNE output. A program based on the Fieldtrip software is created to estimate conditional Granger causality spectra in the time-frequency domain based on these trials.

End-to-end simulations support the correctness of the pipeline with single and multiple subjects. Group-irMxNE makes no attempt to generalize a solution between

subjects with clearly distinct patterns of source connectivity, but shows signs of doing so when subjects’ patterns of activity are similar.

The pipeline is applied to MEG data from the facial emotion protocol in an attempt to validate the Adolphs model. Both irMxNE and Group-irMxNE place numerous sources during post-stimulus periods of high evoked power but neglect those of low power. This identifies a conflict between power-based localizations and information-centric processing models. It is also noted that neural processing is more diffuse than the neatly specified Adolphs model indicates. Individual and group results generally support early processing in the occipital, parietal, and temporal regions, but later stage frontal localizations are missing.

The morphing of individual subjects’ brain topology to a common source-space is currently inoperable in MNE. MEG data is therefore co-registered directly onto an average brain, resulting in loss of accuracy. For this as well as reasons related to uneven power and computational limitations, the early stages of the Adolphs model are only generally validated. Encouraging results indicate that actual non-stationary group connectivity estimates are produced however.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

This work undertakes to develop a methodology and an accompanying software pipeline to perform a group network connectivity analysis on functional neuroimaging subjects. The pipeline is developed for use with magnetoencephalography data, and existing techniques and software are enlisted where possible. Many decisions must be made along the way regarding the suitability of different analysis methodologies, the identification of existing software which implements them, and the extensibility of this software as needed. Instead of being undertaken in the background, these decisions are regarded as an important part of this research. As such, it is hoped that this work not only presents useful results which advance the field, but also assists other researchers in navigating their own work.

The structure of this dissertation is driven by this approach. The introductory chapter provides background material on a broad range of topics which apply to network identification and connectivity analysis, despite not all of these methods being used in this work. The methodology chapter then provides extensive criteria by which a combination of techniques are selected based upon the needs of this work. Modifications and extensions to the selected software are also regarded as part of the methodology, and discussed in this chapter. Results are presented in the next chapter, but detailed interpretation of the results is presented as part of the discussion.

Review of Functional Neuroimaging Technologies

As its name implies, the field of *functional neuroimaging* seeks to observe (or *image*) the functioning of the living brain. There are many neuroimaging technologies, but the mainstream of non-invasive research is currently dominated by two, broadly categorized as either *hemodynamic* or *electromagnetic*. Hemodynamic neuroimaging refers to *functional magnetic resonance imaging (fMRI)*, where neural activity is inferred from changes in blood oxygenation within the brain. Electromagnetic technologies include the related pair of *electroencephalography (EEG)* and *magnetoencephalography (MEG, or MEEG when referring to both)*, which measure the electrical activity of the brain directly based on voltage potentials and magnetic fields respectively. Neither technology is clearly superior to the other, with the two possessing complementary strengths and weaknesses. This work does not endeavor to provide a comprehensive accounting of these technologies, but rather a brief introduction to each, focusing on features related to network identification and connectivity analysis.

Functional Magnetic Resonance Imaging

Magnetic Resonance Imaging (MRI) is a noninvasive neuroimaging technology based on differing magnetic properties at the molecular level within the brain. A *static* or *structural* MRI is a detailed, three-dimensional image of an individual's head and brain. Structural MRIs provide the background image against which functional activation is shown, and are the basis of the computer-generated cortical mesh often used for localization in MEEG. Functional magnetic resonance imaging is based on the differing magnetic properties of oxygenated versus deoxygenated hemoglobin, and neural activity is indirectly inferred based upon the oxygen which it consumes. By taking repeated measurements, it is possible to characterize the activity of the brain in time. For a brief introduction to MRI and fMRI see Noll (2001).

FMRI has the advantage over MEEG that it measures activation directly in space, and the location of the activity under study may be attributed directly to a location within the subject's brain. This avoids the necessity of solving the inverse problem as is required in MEEG. FMRI has the disadvantage of a relatively low temporal sampling rate, limited to large fractions of seconds. Neural activity occurs at the millisecond level, and therefore fMRI cannot fully represent the details of this activity in time. In the context of connectivity and other analyses, this prevents fMRI from accessing the frequency bands at which most neural activity occurs.

In fMRI, neural activity is inferred from oxygen consumption in the brain, but oxygen consumption is only indirectly related to the electrical activity of the neurons via the *hemodynamic response function (HRF)*. The process of estimating electrical activity from oxygen consumption requires a *deconvolution* operation, which depends upon knowing the shape of the HRF. Unfortunately, this shape can only be estimated and, critically, has been shown to differ throughout the brain. This disrupts the order of temporal precedence often used to infer connectivity, and therefore Granger causality is counter-indicated for use with fMRI data (David, et al., 2008).

Electroencephalography

Electroencephalography is among the oldest functional neuroimaging technologies, and is based on measuring electrical potentials (voltages) at the scalp. Scalp potentials are directly caused by neural electrical activity, but location is not measured as it is with fMRI. Historically, many EEG analyses are performed in *sensor-space*, or on the scalp potentials, with no attempt to identify the neural sources of these potentials. For an introduction to EEG, see Baillet, Mosher, and Leahy (2001).

Alternatively, neural activity may be *localized*, or estimates made of the locations and activities of the neural sources responsible for the potentials measured at the scalp. This involves estimating solutions to the *forward* and *inverse* problems, which together link measurements at the scalp to source activity within the brain. (A useful analogy is estimating the number, brightness, and changes over time of a group of light bulbs from the pattern they cast on an opaque shade.) Localization requires detailed information about the particular subject's head and brain anatomy. A structural MRI is desirable for this, although the topology of a standard average brain is often used when one is unavailable. The localization process in EEG is imprecise, and is confounded by electrical conduction effects of the intervening tissue between the neural sources and the scalp.

EEG measurements take place at millisecond resolution and, since neuronal electrical activity is measured directly, there is no analog to the HRF in fMRI. As opposed to fMRI studies, the hallmark of EEG is the detailed analyses of time-related activity in the brain, with localization to specific structures often being of secondary concern. The excellent temporal resolution of EEG makes it easily able to access the frequency bands of most neural activity, but the distortion of timecourses may hinder connectivity analysis.

Magnetoencephalography

Magnetoencephalography is also based directly on neural activity. Instead of electric potentials however, MEG measures the magnetic fields generated by neural electric currents (recall the "right hand rule" from introductory physics). MEG shares the technical properties of EEG, with several key differences. Magnetic fields are unaffected by the tissue of the head, so the smearing effect upon localizations and timecourses is avoided. The measured magnetic fields are weaker than EEG potentials, and evoked response experiments therefore require larger numbers of trials to achieve a similar *signal-to-noise ratio (SNR)*. For an introduction to MEG see Baillet, Mosher, and Leahy (2001). An MEG system and subject are seen in Figure 1.

The MEG sensors are not directly attached to the subject's head as are EEG electrodes and the spatial relationship between the head and the sensors may vary from session-to-session and subject-to-subject. Due to this variable geometry it is uncommon (although not impossible) to do sensor-space analyses with MEG. Most MEG analyses are performed in *source-space*, and claimed localization accuracy is within 5mm of the actual sources. MEG thus offers fMRI levels of spatial accuracy along with EEG levels of temporal resolution. This is a best-case scenario however, and difficult to achieve in practice.

Figure 1: MEG System and Subject

A patient being prepared for scanning in an MEG system. Electrodes on the scalp are used for concurrent EEG recording as well as for the detection of ocular activity (eye blinks) during experimental trials. Magnetic sensors are contained in the inverted bowl-shaped dewar above the subject's head. Not only does the head geometry differ from subject-to-subject, but the relationship between the head and the MEG sensors does as well.

Neural Localization in MEG

Localization is the process of estimating the location and temporal activity of sources within the brain that are responsible for observed measurements at the sensors. Localization is not necessary in fMRI because measurements are taken directly in space. It is also not necessary in EEG when performing sensor-space analyses, but may be performed when a source-space analysis is desired. Localization is performed in almost all MEG experiments.

MEG localization may take place on either a three-dimensional volumetric grid, or the two-dimensional surface of the cortex. Most MEG systems utilize at most several hundred magnetic sensors, but the source-space often contains thousands of potential source locations. As such the solution is underspecified, and restricting the solution to the cortical surface, which contains fewer points than a model of the full brain at the same resolution, is common. This is called a *cortical projection*, and the irMxNE localization algorithm used in this work makes this assumption. Examples of distributed cortical projection localizations may be seen in Figure 4 and Appendix D.

Cortical projection requires a digital model of the surface of the cortex. In computer science a model of a surface is often represented by a *mesh.* This term originates in graph theory, and a cortical mesh is comprised of a set of points on the cortex and a set of undirected edges indicating which points are adjacent to one another. A mesh of the cortical surface may be constructed automatically from a structural MRI using the Freesurfer software (Dale, Fischl, & Sereno, 1999).

The locations of potential sources are usually taken to be the nodes of the cortical mesh. In MEG, sources are considered to be equivalent current dipoles, and the strength of the dipole corresponds to the activity at that location. In addition to strength, dipoles have an orientation (think of a simple bar magnet, with some strength, and also a spatial direction). In a *fixed orientation* solution, each node corresponds to a single dipole with its orientation fixed at the geometric normal to the cortex at that location (i.e. perpendicular). In a *loose orientation* solution there is still a theoretical one-to-one correspondence between nodes and sources, but the orientations of the sources are allowed to vary over the timecourse of the solution. In practice, loose orientations are usually implemented by assigning three sources with fixed orthogonal orientations to each node. Since loose orientations confound the measured value of the sources with their orientation, they are difficult to use in post-localization processing, and this work utilizes fixed orientations exclusively.

The MEG *forward solution* expresses the influence of each potential source upon each sensor in the array. This solution is dictated by the shape and position of the subject's head and the folding of the cortex, and is constrained by the principles of electromagnetic propagation (Gramfort, Strohmeier, Haueisen, Hamalainen, & Kowalski, 2013). The measurements of the sensors each represent a linear combination of the sources' activity, and the forward solution is generated directly by a computerized algorithm. Once constructed, the forward solution may be applied to translate source-space activity into sensor-space.

The MEG *inverse solution* translates from sensor-space to source-space. Despite being constrained to the cortex, there continue to be more sources than sensors, and the solution to the inverse problem remains underdetermined. In other words, there are many patterns of source activation which are able to explain a given pattern of measurements at the sensors. For this reason, additional constraints must be applied to the solution. The inverse problem is usually solved numerically, and the program used to do this is commonly referred to as a *localizer* (because it generates source localizations from sensor-space data) or a *solver* (because it numerically solves the inverse problem).

The process of introducing additional constraints to an underspecified problem is called *regularization*. The numerical solver operates by minimizing the value of a cost function, which contains these regularization terms. The cost function must be able to be expressed in a form for which a known numerical minimization algorithm exists, and this is a very specialized field. A *norm* is an operation which translates a matrix into a scalar value. The relationship between the sensors and sources in MEG is expressed as a function based on matrices (Equation 5), and the cost function being minimized is expressed in terms of norms. Different norms promote different attributes in the solution to the inverse problem. The commonly used distributed localization in the *MNE* software (*Minimum Norm Estimation*) is based upon the L2-norm, which results in the numerical solver producing the solution with the least squared error. The MxNE family of solvers is based on adding additional norms to the cost function in order to produce solutions with useful properties. Figure 4 shows an example of an L2-norm distributed localization.

Two broad categories of localizations are relevant to this work. A *distributed localization* estimates neural timecourses for every source. In other words, the entire source-space is retained, even though some sources may be extremely active while others nearly inactive. Distributed localizations provide no assistance with data reduction in the spatial domain. On the other hand, a *sparse localization* attempts to reduce the number of sources used to explain the sensor-space data, thus assisting with spatial data reduction.

When performing a connectivity analysis, the source-space must be reduced to a manageable size in some way, and sparse localization is a useful method to accomplish this. The recently developed MxNE family of solvers represent sparse localization algorithms.

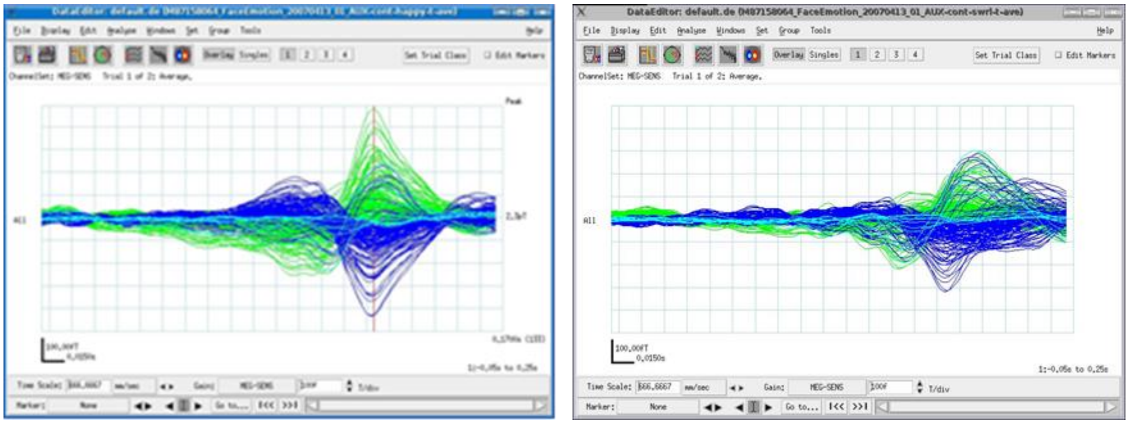
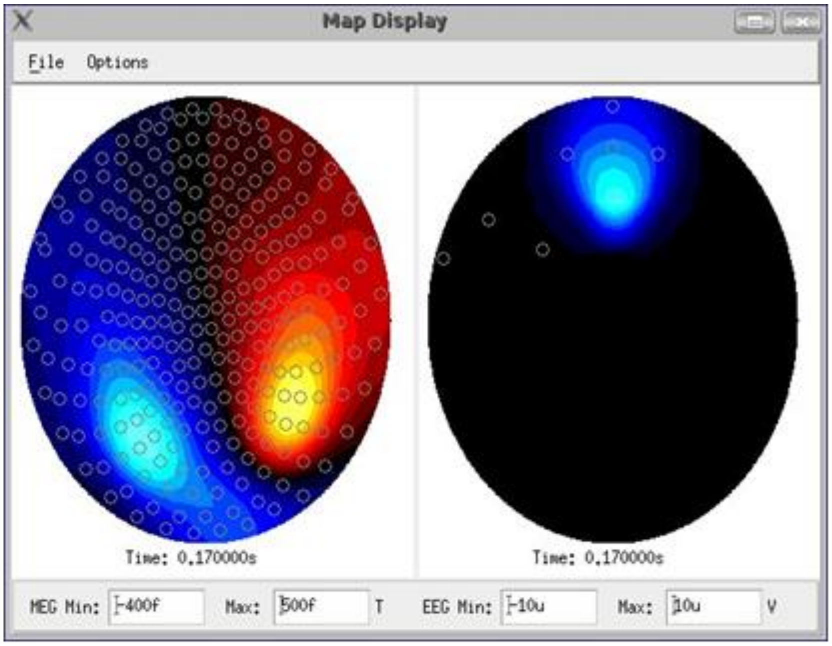


Figure 2: Average Evoked Response Timecourses

Averaged MEG traces for the happy face (left) and swirled face (right) conditions in a single subject. The window shown extends from 50ms pre-stimulus to 250ms post-stimulus in both cases and vertical scaling is the same as well. Seen in these images are the 100ms visual response and a 160ms response including the face-specific fusiform component.



MEG Min: -400f Max: 500f T EEG Min: -10u Max: 10u V

Figure 3: Average Evoked Response in Sensor-Space

A flat map (sensor-space) showing the MEG (left) and EEG (right) sensor activations corresponding to the vertical cursor in the left panel of Figure 2 (170ms post-stimulus after viewing a happy face). The view is from the top of the head looking downwards, the face is towards the top of the image, and the perspective is distorted by flattening. Note the magnetic activation in the occipital and temporal regions in the left panel. The right panel shows the EEG channels, all of which except the vertical eye-blink electrodes on the face and forehead were inactive during this collection. The electrical activation towards the front of the head may reveal that the subject blinked in response to many of the stimulus trials.

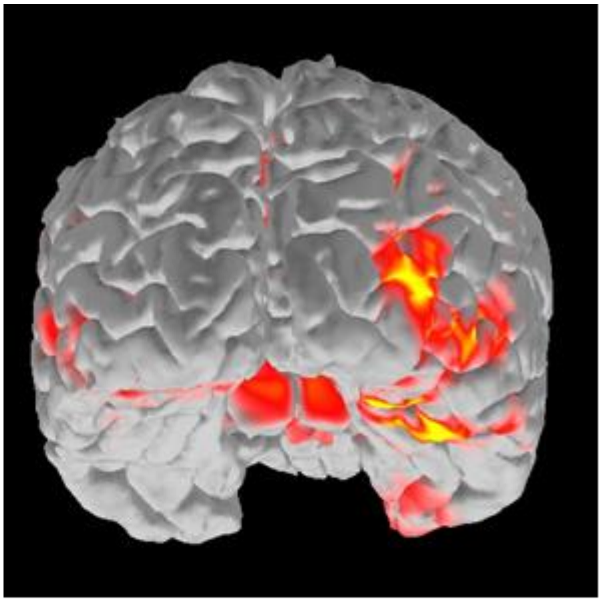


Figure 4: Average Evoked Response in Source-Space.

The cortical projection solution corresponding to the magnetic activation seen in the left panel of Figure 2 and the flat map seen in Figure 3 (i.e. 170ms post-stimulus in the happy-face condition for a single subject). Statistical estimates (dSPM) of cortical activation are seen projected onto the subject's cortical surface mesh. Note activation of the right temporal lobe and fusiform gyrus. This is a 'snapshot' from an MNE 'movie'. The movie itself can be viewed in a jpeg-compatible media player and the cortical activation is then seen to traverse the cortex as post-stimulus time elapses. This is an example of a L2-norm distributed localization, as the full complement of spatial locations is retained. Snapshots from a full movie generated from all twenty-three facial emotion subjects may be seen in Appendix D.

Time Series

A *time series* is a collection of measurements which have a temporal ordering and a fixed interval between them (Schumway & Stoffer, 2006). As opposed to samples in a non-time series experiment, which are presumed to be independent, the values in a time series are assumed to be related to their neighbors. This relationship is formalized by the related *autocorrelation* and *autocovariance functions*, which give the correlation/covariance between values in the series as a function of their distance apart (in samples). The relationship between different time series can be represented in a similar way with the *cross correlation/covariance functions*.

Autoregressive Models

An *autoregressive (AR)* model is a common way to represent a time series in which the present value of the series at any sample is predicted by a linear combination of previous values as well as a random error term. The number of previous values, or *lags*, included in the AR model is called the *model order (p)*. All lags up to p must be included in the model, although the coefficients may be zero-valued (Shumway & Stoffer, 2006). When multiple times series are included in an AR model, and when the value of each is represented by a linear combination of its own past values as well as the past values of the other series, this becomes a *multivariate autoregressive (MVAR)* model. As discussed below, MVAR models form the basis of Granger Causality. Although the random error terms in an MVAR model are often called noise, these terms energize the system and, without them, a stable MVAR system will quiesce to zero. An example of a bivariate MVAR system may be seen in Equations 1 and 2, where the coefficients of the model are given by a_h, b_h, c_h, and d_h, and the noise by ε_0 and ε_1. Coefficients a_h and c_h represent the contributions to the present values of x_0 and x_1 from their own past values. Coefficients b_h give the contributions of the past values of x_1 to the present value of x_0, and d_h, the past values of x_0 to x_1. The latter two sets of coefficients are referred to as *cross terms*, and are central to the conceptual foundation of Granger causality.

$$x_0(t) = \sum_{h=1}^{p} a_h x_0(t-h) + \sum_{h=1}^{p} b_h x_1(t-h) + \varepsilon_0(t) \qquad \textit{(Equation 1)}$$

$$x_1(t) = \sum_{h=1}^{p} c_h x_1(t-h) + \sum_{h=1}^{p} d_h x_0(t-h) + \varepsilon_1(t) \qquad \textit{(Equation 2)}$$

The goodness of fit of AR and MVAR models is estimated based on percentage of variance accounted for. Model fit often continues to improve indefinitely with larger model orders. However, a diminishing benefit is seen with each increase. Therefore, model order is commonly estimated with either the *Bayesian Information Criteria (BIC)* or the *Akaike Information Criteria (AIC)*. By minimizing these criteria, one attempts to identify the optimal model order that compromises between explaining the most variance and using the shortest possible model.

Two important assumptions made by the AR representation of a time series are stationarity and linearity of the system being modeled. Stationarity is discussed in detail below. The residuals are also assumed to be uncorrelated, which is difficult to support when a time series is generated by sampling a physical phenomenon (Friston, Moran, & Seth, 2013).

In signal processing applications, the fitting of AR models is often preceded by some type of frequency domain filtering to remove artifacts and limit the data to a band of interest. It is cautioned that filtering changes the AR structure of the model in unpredictable ways, and that this type of filtering should be done judiciously (Barnett & Seth, 2011). In practice, it is difficult to avoid filtering operations completely in neuroimaging analysis.

The estimation of AR models from empirical data is non-trivial and problematic. Algorithms to accomplish this are an active area of research (Schlogl, 2006). Along with stationarity, the imprecision of AR estimates is a driving force behind the development of nonparametric algorithms (where the model order *p* and the AR coefficients are the parameters to be avoided) (Dhamala, Rangarajan, & Ding, 2008a; Dhamala, Rangarajan, & Ding, 2008b).

Stationarity

Stationarity is a property of time series which requires that the probabilistic behavior of every sub-sequence of values (of any length, at any position) is identical to that of any other sequence. This is called *strong* or *strict* stationarity. This definition is often relaxed in a form called *weak*, *wide-sense*, or *covariance* stationarity with the following conditions:

- The mean value of the series is independent of time
- The autocovariance function of the series is a function of only the lag between values *(h)* and is not a direct function of time (i.e. the relationship of the any value to its own past values remains constant over time, and also note that this is a stronger condition than simple, constant pointwise variance)

Multiple time series may be *jointly* stationary, under similar sets of conditions. Under the multivariate conditions, the cross-covariance is controlled in the same way (Shumway & Stoffer, 2006). In the context of an MVAR model, joint stationarity implies that the structure of the system remains constant over the time covered. A violation of stationarity indicates that this structure may change during the period in question (i.e. a dynamic system).

Tests for stationarity include the *Augmented Dickey-Fuller (ADF)* and *Kwiatkowski–Phillips–Schmidt–Shin (KPSS)* tests. Within an MVAR system these are usually applied to the series individually and it is assumed that if the series are independently stationary that they are jointly stationary as well. Confusingly, both tests require that a window length be specified and advice is given to test multiple size windows. It may then be found that a series is stationary at some lengths but not at others. Even worse, the tests may not come to the same conclusion, and uses are typically counseled to regard this as a non-stationary outcome (Seth, 2010).

Stationarity becomes confusing in a philosophical sense as well. Presumably a signal which contains information is somehow changing in time. How can such a signal meet the definition of stationarity?

In time series analysis, non-stationarity is resolved by a combination of detrending, differencing, and variable transformation. This complicates interpretation of the results, and does not address the issue of a changing structure (Shumway & Stoffer, 2006). Additionally, in signal processing stationarity may be addressed with the application of band-stop filtering (where it is assumed that the non-stationary portion of the signal is band-limited and may thusly be removed). In neuroimaging applications of Granger causality analysis it is reported that this approach causes more problems than it solves and should be used judiciously (Barnett & Seth, 2011).

In neuroimaging the assumptions of piecewise linearity and stationarity are often invoked prior to the use of piecewise analysis, or analysis of segments of a neural response within which these assumptions are thought to be met. This is a valid approach, but segmenting the response into appropriate time windows is difficult and subjective. Additionally, shorter time segments may result in reduced statistical power.

Many neuroimaging analysis methods require some form of stationarity, usually wide-sense, or joint wide sense. Stationarity is poorly understood in the neuroimaging literature. Unfortunately, it is common to apply a method which depends upon stationarity without testing this assumption and understanding its implications. Sample MEG timecourses, which clearly do not meet the definition of stationarity, are shown in Figure 2.

Sampling of Electronic Signals

When a continuous electronic signal is *sampled*, as is done prior to digital processing, it become as time series. The interval between the *samples*, or the *sampling interval*, is commonly written as *dt*. The *sampling frequency* is the reciprocal of the sampling interval, or $f_s=1/dt$.

The *sampling theorem* states that a continuous, band-limited, finite-energy signal may be fully recovered without information loss from a sampled version of itself provided that the sampling frequency is at least twice as high as the highest frequency in the original signal (Haykin, 1989). This theorem provides a critical bridge between analog signals and the ability to process them in the digital domain. If a signal is sampled at less than twice its maximum frequency, a phenomenon called *aliasing* takes place (Haykin, 1989). For this reason signals are often low-pass filtered prior to sampling to guard against this possibility. For a given sampling rate f_s, the *Nyquist frequency*, or $f_s/2$, represents the greatest frequency that is represented in the sampled series (Haykin, 1989).

Spectral Decomposition

Fourier Transform

The *Fourier transform* is a linear, reversible transform of a time domain signal. Formally, it is a *change of basis* to a complex-valued sum of sine and cosine functions of different amplitudes, phases, and frequencies. This is called the *Fourier* or *frequency domain*, and the resulting coefficients denote the parameters of the sinusoidal components that completely reproduce the time signal. Values on the frequency axis are commonly given in units of *Hertz (Hz)*, or *cycles per second*, but *radians per second* is used in more technical works ($w=2\pi f$).

Since the frequency domain coefficients are complex-valued, we have the option of working with *real* and *imaginary* parts, or *magnitude* and *phase*. It is also common to use squared-magnitude, or *power*. We often speak of the *power spectrum*, or a plot of signal power as a function of frequency. This is a handy way to depict the frequency content of a signal. The Fourier power spectrum also represents a variance decomposition of the time domain signal by frequency and, with proper scaling, integrating across frequency recovers the variance of the original time domain signal (Barbour & Parker, 2014). In some formulations this is referred to as *Parseval's power theorem* (Haykin, 1989). Additionally, the power spectrum and the autocovariance function of a wide-sense stationary signal form a Fourier transform pair. This property is referred to as the *Wiener-Khinchin Theorem* (Papoulis, 1991).

Technically, the term *Fourier transform* refers to an integral that is solved in continuous time, transforming a continuous function of time to a continuous function of frequency. This is useful when the analytic form of an input time domain function is known. Tables of well-known *Fourier Transform Pairs* are common in this field (Haykin, 1989). Note that when one transforms to the frequency domain, the time domain is lost. Hence, one works in either the time or frequency domain.

The Fourier transform may also be applied to a discrete or sampled signal. In this form it is called the *Discrete Fourier Transform (DFT)*. A computationally efficient implementation of the DFT that may be applied when the number of samples in the series is a factor of two is the *Fast Fourier Transform (FFT)*. In either case, the input data is formally a time series, and the frequency domain also becomes a series of evenly spaced values in frequency. This method is used on empirical signals where the closed-form is not available. For a brief, accurate, and accessible introduction to the DFT and FFT see Richardson (1978).

The DFT outputs a discrete frequency spectrum with frequencies defined by the sampling frequency and the number of time domain points. If the number of points in the time domain is given by n, and sampling frequency in the time domain by f_s, then the spacing between points in the frequency domain is f_s/n. If the points are indexed starting with '1', then the 'n' frequencies represented in the DFT spectrum are $(n-1)\times(f_s/n)$. This spectrum begins with the zero-frequency *DC term* and continues to just under the sampling frequency (one point short of it). However, the Nyquist frequency (see the sampling theorem above) falls half way along this series of frequencies and the coefficients beyond it are redundant (and mirror those below).

The power spectrum of a series is related by a simple formula to the AR coefficients and the noise variance of the series (Shumway & Stoffer, 2006). The term *parametric spectral estimation* refers to the process of first estimating the AR coefficients of a series (the parameters) and then using these coefficients to generate the power spectrum. For example, for a univariate autoregressive signal $x(t)$:

$$x(t) = \sum_{h=1}^{p} a_h x(t-h) + \varepsilon(t)$$

The Fourier transform is given by (assuming $\varepsilon(t)$ is a white noise process):

$$X(f) = \frac{\sigma_{\varepsilon_0}}{1-\sum_{h=1}^{p} a_h e^{-j2\pi hf}}$$

And the power spectra is given by:

$$S_x(f)=|X(f)|^2 = \frac{\sigma_{\varepsilon_0}^2}{\left|1-\sum_{h=1}^{p} a_h e^{-j2\pi hf}\right|^2} \qquad \textit{(Equation 3)}$$

For multivariate systems this relationship is expressed with matrices (Ding, Chen, & Bressler, 2006). For the bivariate system given in Equations 1 and 2, the coefficients matrix ***A(f)*** is as follows:

$$\boldsymbol{A(f)} = \begin{bmatrix} 1-\sum_{h=1}^{p} a_h e^{-j2\pi hf} & -\sum_{h=1}^{p} b_h e^{-j2\pi hf} \\ -\sum_{h=1}^{p} c_h e^{-j2\pi hf} & 1-\sum_{h=1}^{p} d_h e^{-j2\pi hf} \end{bmatrix}$$

The transfer function ***H(f)*** is the inverse of the coefficients matrix:

$$\boldsymbol{H(f)} = \boldsymbol{A^{-1}(f)}$$

And the noise matrix ***Σ*** is:

$$\boldsymbol{\Sigma} = \begin{bmatrix} var_{\varepsilon_0} & cov_{\varepsilon_0\varepsilon_1} \\ cov_{\varepsilon_1\varepsilon_0} & var_{\varepsilon_1} \end{bmatrix}$$

And finally, the spectral matrix ***S(f)*** is given by (where '*' indicates the matrix conjugate transpose):

$$\boldsymbol{S(f)} = \boldsymbol{H(f)\,\Sigma\, H^*(f)} \qquad \textit{(Equation 4)}$$

The DFT assumes that the sampled time domain signal repeats itself indefinitely in both directions (Haykin, 1989). This is called *windowing*, and refers to the signal only being visible within a specific window in time. The repetition of this window leaves the DFT susceptible to artifacts generated by the windowing processes. These artifacts are often addressed via *padding* the time series with zeroes. Since the FFT algorithm benefits from series of powers of two in length, padding to the next larger power of two is very common (Torrence & Compo, 1998). However, this procedure is also unreliable in that the nearest power of two is an arbitrary distance from the current length. (The wavelet transform has the same issue, and the time series in this work are zero-padded to 2nd-next power of two.)

Achieving smooth, reliable estimates of the Fourier power and cross power spectra from empirical signals is non-trivial. (The intuitive solution of sampling the signal for longer durations simply divides the spectra below the Nyquist frequency into larger numbers of points at finer resolution, but with the individual coefficient estimates continuing to be unreliable.) More effectively, the signal is repeatedly sampled with a reasonable window length, the DFT taken of each and converted to power, and then the power estimates averaged (Dhamala, Rangarajan, & Ding, 2008a; Dhamala, Rangarajan, & Ding, 2008b). When this is not sufficient, additional smoothing may be achieved via the multi-taper method of Thompson (1982), which employs *discrete prolate spheroidal sequences (DPSS)* (Slepian & Pollak, 1961).

Wavelet Transform

The *wavelet transform* was developed in response to the forfeiture of the time axis when working in the Fourier frequency domain. Wavelets preserve time while decomposing the frequency content of a signal. This allows the representation of a signal whose frequency content changes with time. This is called the *time-frequency (TF) or wavelet domain* (Torrence & Compo, 1998). (A windowed version of the DFT called the *Short Time Fourier Transform (STFT)* has been developed in an attempt to preserve the time domain, but has been found inferior to the wavelet transform (Ding, Bressler, Yang, & Liang, 2000).)

A common analogy for the wavelet transform is a sliding window which traverses the duration of a piece of music and responds with a scalar value each time a ‘middle-C’ is played. The scalar value represents the strength of the note. Now imagine one such window for each note or string on an instrument. Using this method it is possible to represent an entire piece of music, with each note being localized in both frequency and time.

This sliding window contains a wavelet function. The *mother wavelet* is the prototype for this function, which is *scaled* to different lengths for each frequency to be measured. The scaled copies are called *daughter wavelets*, and the sliding operation is called *translation* (i.e. translation along the time axis in the Cartesian sense). Different mother wavelet functions are available depending on the properties needed in an application. One of the most common wavelets is the *Morlet wavelet*, which is used in this work (Goupillaud, Grossmann, & Morlet, 1984). Note that lower frequencies have longer wavelengths, and the window for low notes is longer than it is for high notes.

The mother wavelet itself may require parameters to specify its properties. For example, the Morlet wavelet takes a parameter 'w_0', which specifies the number of sinusoidal oscillations within its Gaussian envelope. These parameters are usually set non-empirically according to convention and the judgment of the investigator (the most common value of w_0 is '6' (Torrence & Compo, 1998), although '12' is used in this work based on personal communications with Mukesh Dhamala (October 2013)). The Morlet wavelet is complex-valued, and the resulting wavelet spectrum is also complex and shares many properties with the Fourier spectrum. See Figure 5 for the waveforms contained within the Morlet wavelet.

Like the Fourier transform, the wavelet transform exists in continuous and discrete-time forms. As with the continuous Fourier transform, the *continuous wavelet transform (CWT)* is not suitable for digital processing. Although terminology in the literature is inconsistent, there are two general versions of the discrete wavelet transform, the *discrete wavelet transform (DWT)* proper, and the *continuous time-discrete wavelet transform (CT-DWT)*. Since the number of samples necessary to represent a signal changes with frequency, the length of the output series of the DWT decreases from high to low frequencies. This is useful when trying to avoid redundancy of information (in applications such as signal coding and compression), but effectively compresses the time domain by different factors at different frequencies, making comparison between frequencies difficult (i.e. the wavelet spectra is not square, having fewer points per row at lower frequencies). The CT-DWT, commonly used in research where display of information is of primary concern, allows for redundant information and maintains the same one-to-one correspondence with time at all frequencies. This work deals strictly with the CT-DWT (Torrence & Compo, 1998).

While the DFT generates a frequency axis with fixed values determined by the sampling frequency and number of points in the time series, the user of the wavelet transform may define the frequencies which they desire to measure. The convention, most likely originating with the DWT proper, is to use the minimum non-zero Fourier frequency and then increase by factors of two (Torrence & Compo, 1998), which forms a complete basis set without replicating information. This work uses the same frequencies as would be contained in the equivalent Fourier spectrum for ease of comparison, which results in additional redundant information being generated. (Frequencies above 60 Hz, which this work removes during preprocessing, are not displayed in the spectra.)

Daughter wavelets are created by multiplying the time axis of the mother wavelet by a *scale* corresponding to each desired frequency. This scale is directly related to frequency by a formula specific to the mother wavelet selected (Torrence & Compo, 1998). Because the DC, or zero-frequency term of a series is invariant in time, it cannot be estimated by the wavelet transform (the daughter wavelet would have infinite length). However, the series mean may be substituted here. While the DFT simply converts an 'n'-point time series to an 'n'-point frequency spectrum, the wavelet transform generates a 'p' × 'n' matrix, where 'p' is the number of requested frequencies, and 'n' is the 'n'-points from the original time series.

Unlike the DFT, the peak amplitude of the wavelet spectrum does not remain constant across frequencies. Rather, in order to maintain conservation of power across frequency, the response, which necessarily becomes wider, also becomes shorter at lower frequencies. The user must choose between constant power and constant amplitude. Conservation of power is not a priority in this work, but differing peak amplitudes complicate interpretation, and the latter option is chosen. This is done by applying the amplitude corrections of Liu et al. and Veleda et al. to the power and cross power spectra respectively (Liu, Liang, & Weisberg, 2007; Veleda, Montagne, & Araujo, 2012). See Figure 6 for an example of this.

The wavelet transform encounters problems with *edge effects* at the ends of the time series being transformed. These effects occur when the time-scaled daughter wavelet begins to fall off the end of the series (whereupon the wavelet may either wrap to the other end of the series or enter into a region of zero-padding – this issue is shared with the DFT). Torrence and Compo formally define a *cone of influence (COI),* outside of which the value of the wavelet spectra cannot be interpreted reliably because of these edge effects (1998). This segment gets longer with lower frequencies, and the lowest frequencies are contained within this region entirely (with the minimum frequency depending on the total length of the time series). For an example of the COI, see Figure

15 and subsequent wavelet power spectra where it is depicted using white lines. Note, however, that the value of the wavelet spectra within the COI is only suspect in an absolute sense (e.g. assuming zero padding, the value of the spectra in this region is ‘diluted’ by the presence of the zeros within the wavelet window). Spectral values within the COI may still be compared to values at the same location in other spectra (since the ‘dilution’ is equal in both cases). This becomes important when performing inference and comparing against a null spectra, and in the measures of Granger influence, which are a ratio of two spectra. For reference, this work displays the COI on all spectra, although it technically only applies to the values of the raw wavelet spectra (versus Granger spectra and inference spectra).

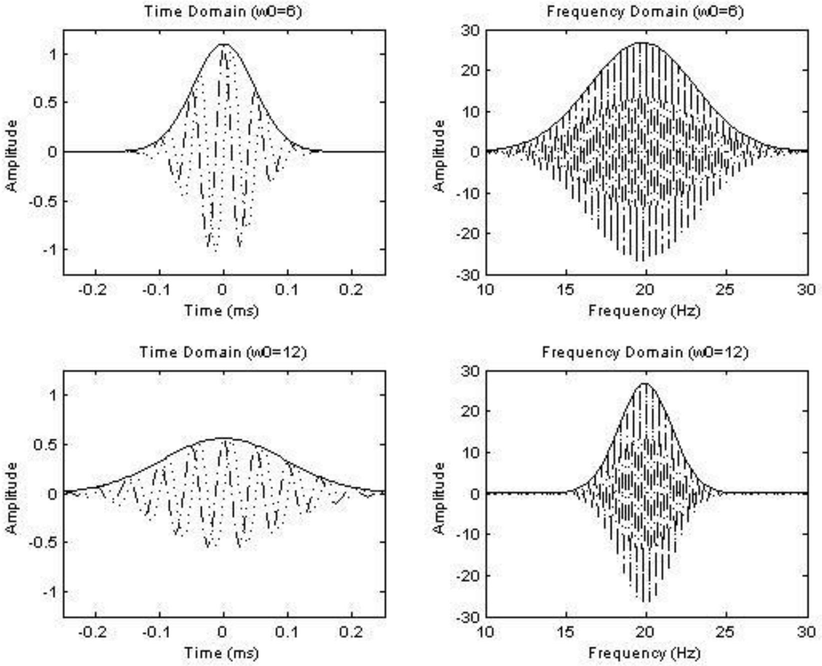


Figure 5: Morlet Wavelets in the Time and Frequency Domains

Morlet wavelets in the time (left) and Fourier frequency (right) domains. The Morlet wavelet is complex-valued, and the real (dashed line), imaginary (dotted line), and absolute value (solid line) are shown in the figure. In essence, the Morlet wavelet is a complex-valued sinusoid within a Gaussian envelope. As is visible on the right, the wavelets shown here are scaled to select components at 20 Hz. Wavelets for parameter *w0=6* (top) and *12* (bottom) are shown, where the parameter represents a tradeoff between resolution in the time versus frequency domains. Note that both Gaussian waveforms and sinusoids are their own Fourier transforms, so that the time and frequency figures here appear very similar.

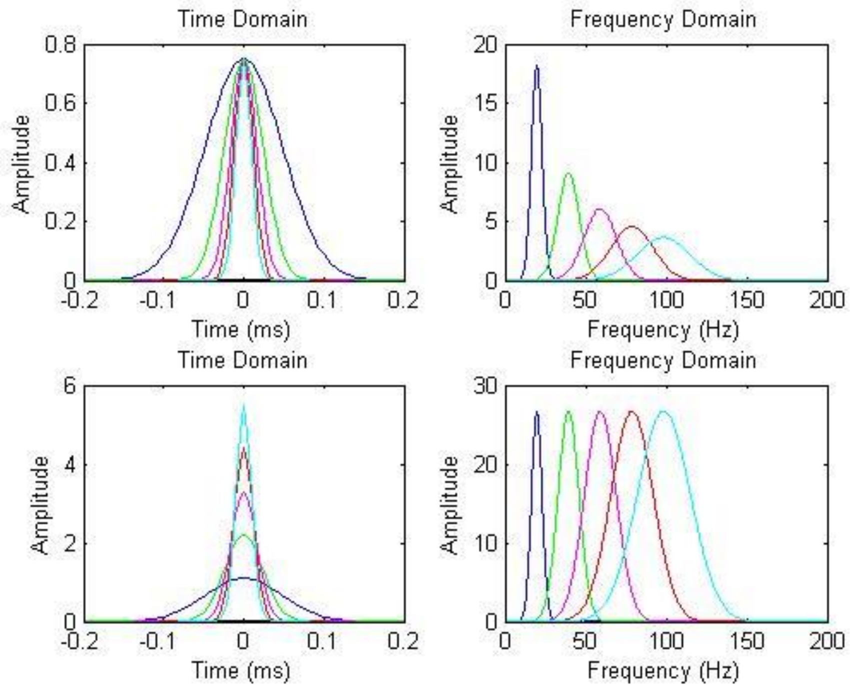


Figure 6: Morlet Wavelets and Amplitude Scaling

Absolute values of Morlet wavelets scaled for 20 (black), 40 (green), 60 (magenta), 80 (red), and 100 Hz (cyan) in the time (left) and Fourier frequency (right) domains. The raw wavelet envelopes in the upper right pane are scaled such the area under each is constant, representing a conservation of power in the frequency domain, but which is confusing because peak amplitude is reduced as the lower frequency wavelets become wider in the time domain (left). Scaling is applied in the lower panes such that power is no longer conserved but peak frequency remains constant across the range (Liu, Liang, & Weisberg, 2007; Veleda, Montagne, & Araujo, 2012). Since the wavelet transform is equivalent to pointwise multiplication in the Fourier domain, the envelopes in the right hand panes may be thought of as selecting Fourier frequencies. As the Fourier frequency rises the range of frequencies selected becomes wider.

Cross Power Spectrum

The power spectrum is the squared-magnitude of the complex-valued raw spectral output from the DFT or CT-DWT (equivalently the product of each complex value in the spectra with its own complex conjugate). As such, the power spectrum is real-valued. When spectral decomposition is performed on multiple time series, there also exist pairwise cross power spectra between the signals. The cross power spectrum is defined as the product of the values of one raw spectrum (complex) with the complex conjugate of the other raw spectrum (complex), and the result is complex-valued. Note that there are actually two cross power spectra, depending on which of the original values is conjugated. These two cross power spectra are complex conjugates of one another. (This property becomes important in the next section.)

Since the cross power spectra are products of two raw spectra, *all cross spectra are power spectra*. Terminology in the literature is inconsistent, but this work will use the term *cross power spectra*. The cross spectrum represents the power transfer between the signals decomposed by frequency. If the signals are independent, the cross power spectrum is zero at all frequencies. This cross power represents the sum of the power passing in both directions between the series. Decomposing this power by direction is of interest because it represents the direction of influence in the system being modeled.

The Fourier cross power spectrum is also directly related to the MVAR coefficients as noted above (Equation 2).

Spectral Matrices

The spectra resulting from the Fourier and wavelet transforms have a number of unique properties that become useful with regard to the Wilson spectral decomposition algorithm. When considered together, the power spectra and cross power spectra of a system form a spectral matrix or, in Wilson's terminology, a *matrical function* (1972; 1978).

In our context, a matrical function is simply the power and cross power spectra of a system arranged in a three dimensional matrix. The dimensions of this matrix are 's' x 's' x 'n', where 's' is the number of sources in the system, and 'n' the number of time points in the sampled series (assuming the wavelet transform uses the 'n' Fourier frequencies). For a bivariate system, where s=2, there are two power spectra (for series '1' and '2') and two cross power spectra (between series '1' and '2' and '2' and '1', which are complex conjugates of one another). Within each 's' x 's' matrix, the diagonal elements represent the power spectra, and the off-diagonal the cross power spectra.

At each of the 'n' points, the power spectra on the diagonal are real valued, and the cross power spectra in the off diagonal locations are complex, and conjugate symmetric of one another. Combined, these properties mean that each 's' x 's' matrix is *Hermitian*.

Connectivity Analysis

As mentioned previously, neuroimaging has long been concerned with identifying locations of neural activity associated with different tasks and cognitive processes. The details of this activity, such as its timing and intensity, have also been of interest. More recently, and corresponding to the view of the brain as a collection of networks (Rubinov & Sporns, 2010), interest has grown around measuring connectivity between active regions. Although this work hoped from the outset to utilize Granger causality analysis, a broad view of connectivity is presented below.

Friston defines three types of connectivity of interest in neuroimaging (Friston, 2011). The first, *structural connectivity*, refers to the anatomical connections within the brain itself, and is a useful concept because the structure of connections in the brain is thought to constrain other forms of connectivity, but is not sufficient to predict them (Bassett & Lynall, 2013). A useful analogy is that of a telephone network, where the configuration of switches and wiring constrains the calls that may be placed, but does not predict who will call whom.

Functional connectivity is a statistical connectivity relationship between measurements of neural activity, but one which says nothing about the model of neural activity which generates it (Friston, 2011). In this case the measured data may indicate a relationship between neural locations, but no model of these sources and their relationship is required. Many exploratory techniques, including Granger causality as applied in this work, fall into this category.

Effective connectivity refers to an explicit relationship between neural regions based on a mathematically defined generative model of their activity and the connection between them (Friston, 2011). This is the most rigorous method of connectivity analysis, but the details of neural function required render it more suitable for confirmatory research. *Dynamic Causal Modeling (DCM)* is the primary example of effective connectivity (Friston, Moran, & Seth, 2013).

A pair of non-exclusive terms regarding connectivity are also relevant to this work.

Dynamic connectivity acknowledges that the pattern of connectivity within a system changes with time. Many early analyses assumed a static network configuration, which is not supported. Techniques which do not rely on the assumption of stationarity are required to assess this (Hutchison, et al., 2013). As seen in this work, connectivity estimation techniques may come to depend upon stationarity due to the details of their implementation versus their original formulation. Granger causality is one such technique where, when based upon MVAR estimation, it depends upon stationarity, but a non-parametric version which does not is also available.

Directed connectivity acknowledges that influence between regions is not reciprocal, but may be unidirectional, or bidirectional to unequal degrees. Directed connectivity requires a technique which is able to decompose influence by direction. According to the review of Collura, non-directed connectivity estimation techniques include *coherence, phase similarity, spectral correlation coefficients, comodulation, asymmetry*, and the sum or difference between channels. Directed connectivity techniques include *Granger causality, dynamic causal modeling, partial directed coherence*, and the *dynamic transfer function* (2008).

Granger Causality

Granger Causality is a versatile technique for exploratory analysis of connectivity between time series (Granger, 1969). Originally formulated based upon MVAR modeling of systems of time series, Granger causality is based upon temporal precedence. Within MVAR models, the present value of each series is predicted based upon that series' past values as well as the past values of the other series (see Equations 1 and 2). If prediction of one series' current values is improved by the inclusion of another series' past values as well as its own, then the second series is determined to cause the first. Inference is accomplished via an *F*-test on the residual variances of the two models. This conclusion depends upon all relevant variables being included in the analysis, and the literature has adopted the term *Granger causality* to denote causality under this assumption. Mistaken conclusions of causality are possible when series not included in the analysis are influencing series that are. A very accessible introduction to the mechanics of Granger causality is provided by Ding, Chen, and Bressler (2006).

Although not usually referred to as such, Granger causality based upon MVAR modeling is *parametric*, being based upon estimation of the regression coefficients and model length. When based upon MVAR estimation, Granger causality inherits the assumption of stationarity from this procedure.

Frequency Domain Granger Causality

In addition to the time-domain formulation above, Granger's original work outlines a frequency-domain framework as well (1969). The only difference is that causality is decomposed by frequency, and the specificity gained by doing so makes frequency domain Granger causality a popular choice in research. In order to calculate Granger causality in the frequency domain, an additive decomposition of the cross-power spectra is required. This may be accomplished in two ways.

MVAR Decomposition

The easy availability of the cross terms in the MVAR model make the directional, additive decomposition of influence between series a simple matter. This extends to the frequency domain, where the estimated MVAR cross-coefficients are used to generate the directional decomposition of the cross spectra. As mentioned previously, stationarity is assumed.

Wilson Decomposition

In order to avoid MVAR estimation and its accompanying assumptions, an alternative method of decomposing the cross-power spectra is needed. This would allow the spectra output by the Fourier transform to be used directly. Fortunately, a numerical algorithm developed by Wilson and based upon the Hermitian nature of spectral matrices may be used here (1972). Happily, this algorithm is guaranteed to converge to a deterministic solution (Wilson, 1978).

When the Wilson algorithm is used to decompose the cross-power spectra by direction, the need for MVAR estimation is eliminated and the ensuing assumption of stationarity is avoided. In order to represent the now potentially dynamic influences between series, it is appropriate to use the wavelet, or time frequency domain in place of the Fourier domain. This, in turn, requires the Wilson algorithm to operate on wavelet instead of Fourier spectra. The prerequisite conditions set out by Wilson appear to allow this (1972), and the application of the Wilson algorithm to wavelet spectra is the basis of the non-parametric Granger causality works of Dhamala, Rangarajan, and Ding (2008a; 2008b). (However, also note the discussion of a skew towards lower peak frequency when applied to wavelet but not Fourier spectra in Appendix E.)

The cross-power spectra, as decomposed by the Wilson algorithm, takes the place of that estimated from the MVAR coefficients (at least in the Fourier domain, as there is no equivalent to the MVAR representation in the wavelet domain), and the calculation of Granger causality estimates proceeds in the same way. Interestingly, much of the literature on Granger causality assumes that the non-parametric form implemented in this way is somehow an inferior estimate of that which would be obtained using MVAR methods. In fact, the situation is reversed, and the MVAR coefficients are used to estimate the spectra instead of generating it directly. Along a similar vein, Barnett and Seth caution against the use of non-parametric Granger causality based upon the computational burden of the numerical algorithm (valid), and upon imprecision of the results, but make no mention of non-stationary data (2014b).

Conditional Granger Causality

To this point in our discussion we have implicitly assumed a pairwise analysis of connectivity between series. Systems of multiple series may be decomposed and addressed in a similar, pairwise fashion, but with inherent limitations. In particular, pairwise results are unable to distinguish between direct causation between series, and causation that is mediated by an additional variable. For example, in a trivariate system with causality $x_1 \rightarrow x_2 \rightarrow x_3$, pairwise analysis will return this pattern of connectivity, but will also erroneously find causation $x_1 \rightarrow x_3$ (although presumably at longer lags). Granger referred to this as a *prima facie* error, and formulated a conditional procedure to address this problem (1969). Referred to in the literature as *conditional Granger causality*, this procedure is considerably more complex, but effectively resolves this issue (Ding, Chen, & Bressler, 2006). Conditional analysis exists in both the time and frequency domains. When analyzing complex system with multiple series, unless there is reason to believe that no opportunity for prima facie errors exists in the data, conditional Granger causality should be used whenever possible.

Measures of Influence

Granger and Geweke both develop power spectral decompositions of the influence between series (Granger, 1969; Geweke, 1982). Both authors develop corresponding *measures of influence* to accompany this. These measures of influence are simply functions of frequency which quantify the directional degree of influence between series in a way which is free of the units of spectral power. These measures are easier to interpret and facilitate comparison between works.

Granger's measures are two values, C_{xy} and C_{yx}, which represent the bidirectional degree of influence of two series upon one another. These values are simply the squared directional portions of the cross power spectra divided by the product of the two individual spectra (a normalization process very similar to the way in which covariance becomes Pearson correlation) (Granger, 1969). As such they are bounded to the range [0-1] and may be interpreted as magnitude squared coherence (Collura, 2008). These measures are used to report results in this work.

Geweke's measures, I_{12} and I_{21}, quantify the causal influence of series upon one another via his intrinsic/causal decomposition of each series. Each measure is a function of frequency defined as the log ratio of the total power of each series divided by the intrinsic power of that series – in other words, the total power in the series divided by the power derived from its own past (Geweke, 1982). These measures are always non-negative but are unbounded. Both Granger's and Geweke's measures are zero when no directional influence exists between series.

By using these normalized measures, the investigator loses track of the relative power of the series. In our case, influence between regions of relatively low neural activation appear comparable to those of high activation. (This property allows averaging over individual causality spectra in the group model developed below, and also renders the causality spectra immune to the reduced wavelet power within the COI.)

Components Analysis

Although not ultimately utilized in this work, components analysis procedures have a number of properties which make them appealing in the context of sparse localizations, network identification, and group solutions. In fact, the only group solution currently known to exist in neuroimaging is based on components analysis (Calhoun, Adali, Pearlson, & Pekar, 2001). For this reason, a discussion of components analysis is presented here.

Components analysis is a statistical procedure which identifies hidden or *latent variables* underlying a set of *measured variables*. This is useful for several reasons:

- Parsimony/Interpretability: The measured variables may be mixtures of a set of latent variables, and thus components analysis reveals the underlying structure of the data more clearly.
- Data Reduction: There are often fewer latent variables than measured variables, and components analysis reduces the dimensionality of the data while retaining the most important aspects of it.
- Error Rate Control: A related benefit of dimensionality reduction is a reduced number of inferential tests required during analysis. When using Bonferroni-style corrections, performing as few tests as possible preserves statistical power.
- Uncorrelated/Independent Variables: Statistical testing and other procedures may require that variables be uncorrelated with, or independent of each other (e.g. multicollinearity of regressors). When measured variables share a common latent variable this assumption is violated. The variables output from components analysis are uncorrelated (PCA) or independent (ICA).

Components analysis has a long tradition in psychology, and most psychologists are familiar with *Factor Analysis (FA)*. The *Principal Components Analysis (PCA)* procedure used in neuroimaging and other fields is directly related to factor analysis. Another common components analysis algorithm used in neuroimaging is *Independent Components Analysis (ICA)*, which was developed in the field of signal processing.

The components analysis algorithms discussed here make several assumptions:

- Linear mixing: the measured variables must be linear combinations of the latent variables.
- Instantaneous mixing: at any instant in time, the measured variables are a function of the current values of the latent variables alone.
- More observations than variables: the number of latent variables returned is capped by the smaller of the number of measured variables and the number of observations minus one.

Formally, components analysis algorithms operate by axis rotation. The input data is represented as a scatterplot in *N*-dimensional space and the axes of this space are rotated to minimize or maximize some criteria. For an accessible, general introduction to components analysis, see Stevens' discussion of Factor Analysis (2009).

Components Analysis in Signal Processing

The field of signal processing is also interested in the identification of latent variables in systems. This is called *blind source separation*, or the identification and separation of multiple signals that have become mixed. An example of this is the *cocktail party problem*, where numerous, independent conversations are taking place in a room and are recorded by several microphones. Source separation refers to the problem of extracting individual voices from the mixed recordings. Signal processing terminology differs from that of factor analysis. In this case the pointwise samples of the recorded signals are called *measurements* and the latent variables are called *sources*. (Note the distinction between latent sources, and the source locations at the nodes of the cortical mesh as discussed previously. In a factor analysis solution, the former become weighted sums of the latter.)

Components analysis is often used in neuroimaging in combination with distributed localizations. A distributed localization generates estimated neural timecourses for each potential source location in the analysis. However, it is likely that only a subset of these locations are actively involved in a neural response. Additionally, when neural activity has some spatial extent, groups of locations may be acting in unison.

When applied to the combined time series of a distributed localization, the observations are n-tuples of measurements at each point in time (with ‘n’ being the number of potential source locations). Thus the components represent spatial groups of locations which tend to behave the same way over time. The spatial morphology of each factor is defined by the loadings of the individual locations onto the factors. The factor scores are often called factor timecourses, and represent the behavior of each source over time. (This is an example of temporal components analysis.)

The application of components analysis to time series data is no different from the application to experimental data in non-time series data sets. In particular, when applied to time series data, the components analysis algorithms are unaware of any temporal ordering of the observations, and the time points are simply treated as individual data points.

Principal Components Analysis

Principal Components Analysis (PCA) was developed alongside factor analysis in the psychological tradition. Both PCA and FA operate on the correlation or covariance matrices of the measured variables. The main difference is that factor analysis inserts estimates of communality on the diagonal of this matrix, while PCA retains the original correlation or covariance estimates. For this reason, PCA components are generally easier to interpret.

The raw PCA output represents an orthogonal set of axes of the same dimension as the number of input variables. The PCA axes, or components, are ordered by decreasing variance accounted for. If the researcher takes advantage of this ordering, it amounts to the assumption that the components of greatest variance relate to the phenomena of most importance in the results (Stevens, 2009). Data reduction is accomplished by retaining those components with large variance and discarding those without.

By itself PCA delivers an orthogonal rotation of the original axes. However, it is commonly followed by Varimax (orthogonal) or Promax (oblique) rotations in order to generate more interpretable solutions. The variance-ordering of the output components mentioned above no longer holds following any rotation, and the variance partitioning is no longer valid following oblique rotation. See Dien for a description of these rotations (2010). It is the recommendation of Dien that Promax be used in all cases (1998). Finally, PCA is subject to misallocation of variance when applied to ERPs (Dien, 1998).

Independent Components Analysis

Independent components analysis was initially developed in signal processing to address the blind source separation problem (Bell & Sejnowski, 1995). Neuroimaging analysis has been one of the primary applications of ICA, and early applications were to EEG ERP data (Makeig, Bell, Jung, & Sejnowski, 1996; Makeig, Jung, Bell, Ghahremani, & Sejnowski, 1997).

ICA operates by maximizing the independence of sources according to some higher-order statistic such as kurtosis. The *central limit theorem* states that the sum or difference of several random variables tends toward a normal distribution regardless of the shapes of the individual distributions of the components. ICA is effectively this process operating in reverse, where the higher-order statistic is used to judge the non-normality of the individual components. When non-normality is maximized, it is assumed that the original contributors to the sum have been identified.

Many different ICA algorithms have been developed, all operating somewhat differently and with slightly different sets of assumptions. Several comparisons taken together appear to favor *Infomax* and *JADE* (Correa, Adali, & Calhoun, 2007; Delorme, Palmer, Oostenveld, Onton, & Makeig, 2007; Lee, 2000; Cardoso, 1999). ICA generally assumes that the latent sources are independent and have non-normal distributions. ICA has several limitations that render it more complex to use than PCA. The amplitude scaling of the resulting factors is inaccurate by some unknown multiplier which must be estimated. The number of latent sources must also be estimated, because this is an input parameter to the ICA algorithm and results differ depending upon it. Finally, most ICA algorithms are implemented via non-deterministic numerical solutions, and therefore the results differ between runs (JADE is an exception). *Multi-start* procedures are an inconvenient solution to this problem (Himberg & Hyvarinen, 2003).

Relationship between Principal and Independent Components Analysis

The neuroimaging literature commonly regards PCA and ICA as exclusive and unrelated to one another. However, this is not necessarily the case, as most ICA algorithms first perform a PCA rotation on the data before submitting it to ICA. This is done in order to estimate the number of latent sources contained in the data and to prewhiten the input data which improves the convergence of the ICA algorithm. Additionally, data reduction may be accomplished via PCA in order to reduce the computing requirements of the numerical ICA solution (Calhoun V. D., Adali, Pearlson, & Pekar, 2001).

Components Analysis, Connectivity Analysis, and Stationarity

The brief reviews presented above are not intended to be a definitive discussion of components analysis. However, they are intended to provide the background needed to address the question of suitability of these popular procedures to connectivity analysis research. Referencing the previous development of MVAR systems, it is seen that, when influence between series is present in a system, the values of the series may be expected to be correlated (by virtue of their common past values). PCA components, on the other hand, are uncorrelated by definition. Indeed, PCA is also commonly used as a whitening procedure for the purpose of removing the correlations between variables. For this reason it seems apparent that the use of PCA as a data reduction scheme is antithetical to connectivity analysis. By a similar token, ICA explicitly relies on the assumption of independent sources, and the behavior of its independence-maximizing axis rotation algorithm is difficult to predict when the sources have a level of dependence. For this reason ICA appears to be counter-indicated as well.

PCA and ICA both create linear combinations, or components, of the sources in a distributed localization. These components represent a best fit to the observations compiled over time. As discussed previously with respect to MVAR estimation, there is no mechanism for these linear combinations to represent a changing system. Therefore it is also concluded that neither PCA nor ICA are suitable for use with dynamic systems.

Group Analysis

Neuroimaging experiments commonly utilize multiple subjects for the same reasons as other scientific work, namely to increase statistical power and external validity. Nonetheless, most neuroimaging analysis methods apply to single subjects, and group methods are exceedingly rare.

Group solutions represent a best fit to multiple subjects, and are useful for identifying commonalities in complex results. With regard to the present work, consider the case where two subjects, analyzed individually, form brain networks with different numbers and locations of sources during their experimental responses. In this case it is not possible to make a direct comparison of these subjects without somehow reconciling their differing networks. However, if a single brain network were generated for both subjects simultaneously, a direct comparison becomes feasible. A group solution between subjects is loosely analogous to an average, and the standard criticisms apply. The single solution hides individual variation, and is expected to fit each subject less well than an individualized result.

Within a neuroimaging analysis, several opportunities exist which enable possible group solutions. If the final results for the individual subjects are comparable, they may simply be combined. An example of this is the average L2-norm distributed solution supported by the MNE software (for an example, see Appendix D). As long as all subjects' results are generated on the same cortical model, their individual distributed localizations may be averaged. Unfortunately, this distributed localization does not meet the needs of the present work (although it is used for weighting by irMxNE).

An opportunity to accomplish a group solution with a limited number of sources exists when components analysis is being employed. In this case, the components procedure may be applied across multiple subjects simultaneously. An example of this approach that is widely applied to fMRI data is that of Calhoun, et al. (Calhoun V. D., Adali, Pearlson, & Pekar, 2001; Eichele, Rachakonda, Brakedal, Eikeland, & Calhoun, 2011). Unfortunately this solution does not meet the requirement of the current analysis (see discussion of components analysis above).

The recently developed MxNE family of sparse localizers as implemented in MNE-Python has many properties useful in this work, including their explicit development for use with non-stationary data. In such a case, and provided the algorithm is amenable, a group solution may be undertaken. This work accomplishes a group implementation of the irMxNE localizer. See the methods section for a discussion of how this decision is made and how the group extension is implemented.

Finally, a concern arises in group analyses where subjects are observed to differ with regard to total neural power generation. Some type of between-subjects power normalization may be needed in order to avoid subjects with the greatest power dominating the group solution.

Brain Networks

A *network* is an abstract concept borrowed from graph theory, being comprised of *nodes* and *edges*. Edges may or may not be directed and, when edges are directed, they may be uni- or bi-directional with different levels of influence in each direction. In neuroimaging, nodes represent regions of neural activity, and the edges connections between these regions. Connections ostensibly represent the flow of information between processing centers, and are usually considered to be asymmetrically bi-directional (Bassett & Lynall, 2013; Park & Friston, 2013; Rubinov & Sporns, 2010).

The brain is organized as a network of networks, where local, specialized units are integrated at a larger scale to support complex cognitive functions. *"The function of a module is to integrate and contextualize the more specialized functions of its* *submodules."* (Park & Friston, 2013). In graph theory this is referred to as a *small-world* or *rich-club* topography, with the hallmark being that only a few nodes at a high level of functionality are involved in the response at the highest level (Park & Friston, 2013).

Many connectivity estimation techniques assume a static network configuration and fixed influence between nodes. This is a useful simplifying assumption because it allows for analysis methods which collapse the time domain, but unfortunately also require stationarity. However, the field has recently begun to acknowledge that these networks are dynamic (Hutchison, et al., 2013), and that neural signals are non-stationarity (Kaplan, Fingelkurts, Fingelkurts, Borisov, & Darkhovsky, 2005; Klonowski, 2009).

The Adolphs facial processing model seen in Figure 8 is an example of a dynamic network (Adolphs, 2002a; Adophs, 2002b). Reciprocal activation, exemplified here where the amygdala activates the striate cortex early in processing, and which is then reversed in later stages, is common in brain networks.

Faces and Facial Emotion

Faces and facial expression indicate many things, including gender, age, race, health status, and psychological state. Darwin proposed that facial expressions have evolved to nonverbally communicate emotional states between individuals (Darwin, 1965). Ekman proposes six universal, basic emotions (happiness, sadness, anger, fear, surprise, and disgust), with a unique facial expression corresponding to each (1972). The emotion of contempt has been suggested as well (Ekman & Freisen, 1986). These expressions are thought to be universal and independent of culture (Ekman, et al., 1987). A standardized set of images of emotional faces are utilized in many facial processing experiments (Ekman & Friesen, 1976).

Development of neural models for the perception of faces is an ongoing effort. This work is particularly interested in those which include not only a functional decomposition, but neural locations as well. One of the first of these is the model of Haxby, Hoffman and Gobbini, seen in Figure 7 (2000). In this model, three bilateral structures make up the core system, the *inferior occipital gyri*, the *superior temporal sulcus (STS)*, and the *lateral fusiform gyrus (LFG)*. The core structures communicate bidirectionally with each other. The STS and LFG then project to a number of regions, including the *intraparietal sulcus*, *auditory cortex*, *amygdala*, *insula*, and limbic system, and the anterior temporal regions. These projections are also bidirectional. Unfortunately this model does not include estimates of timing.

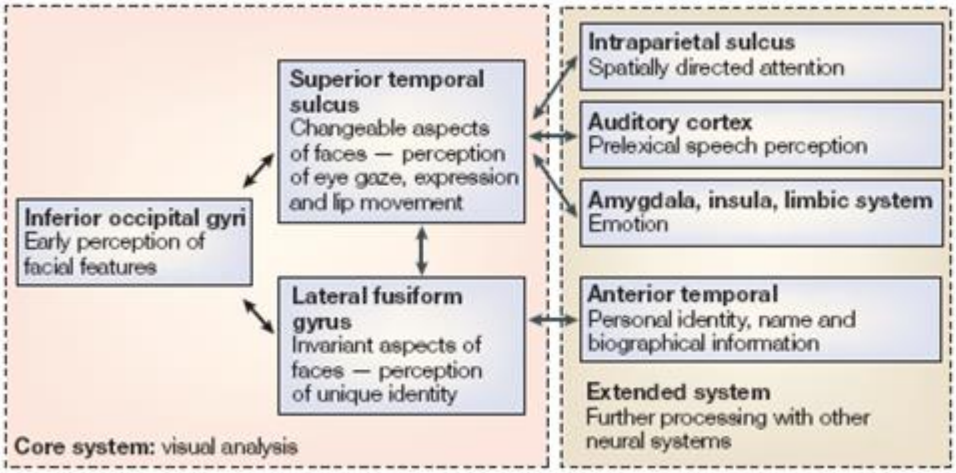


Figure 7: Haxby, Hoffman, and Gobbini Model

The Haxby, Hoffman, and Gobbini model of facial perception (Figure from Calder and Young (2005)).

Among the most ambitious efforts is the model of Adolphs, seen in Figure 8 (2002a; 2002b). In the first stage, activation begins in the midbrain structures, and then enters the amygdala and striate cortex. Stage two involves forward projection from the striate to the STS and *fusiform face area* (FFA). Further projection activates the orbitofrontal cortex, and the amygdala is reactivated. In the final stage the orbitofrontal cortex and amygdala project back to the striate region, the somatosensory cortex, and the limbic system. The first stage of activation takes place in the range 0-120ms, the second 120-300ms, and the third beyond 300ms.

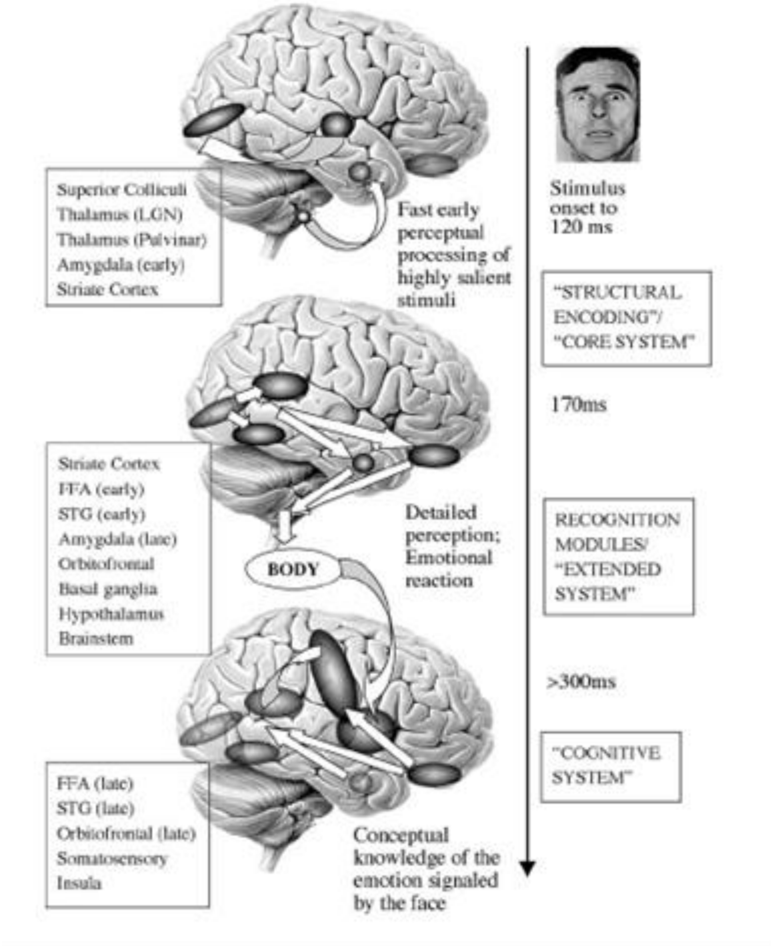


Figure 8: Adolphs Model

The Adolphs model, which includes detailed estimates of timing. (Figure from Adolphs (2002b)).

The neuroimaging analysis pipeline developed in this work is applied to a group of MEG subjects scanned during a facial emotion recognition protocol. The facial perception model is well-researched and offers a good baseline to which the results may be compared, but is also a complex model for the first application of this pipeline. Results produced here advance the fields of MEG analysis and facial perception. Facial perception has implications for many psychological disorders with social and affective components, and contributions to the field may be hoped to assist with the understanding of these problems.

Aims and Goals of this Research

Goal: It will be possible to create a software pipeline to perform a group connectivity analysis with MEG data. Existing software components will be used wherever possible, and will be linked together with new algorithms as needed. This pipeline will not be hindered by the assumption of stationary neural systems. Unless counter-indicated, the first choice for connectivity analysis in this pipeline will be Granger causality. The pipeline will attempt to minimize the specialist knowledge required of the end user and to be as turn-key as possible.

Goal: Appropriate simulations will support the correct functionality of the software pipeline. Simulations will also be provided to assist the reader in understanding this work and interpreting its results.

Goal: Group MEG localizations will be seen for the facial emotion protocol which approximate those of the Haxby, Hoffman, & Gobbini (2000) and Adolphs models (2002a; 2002b). This will be particularly true for cortical regions which are activated early in the model: the striate cortex, fusiform gyrus, and superior temporal sulcus. It is hoped that later, more frontal activations will be seen as well.

Goal: Activation for these regions will approximate that of the Adolphs model, with the striate cortex activating first followed by the fusiform gyrus and superior temporal sulcus.

Goal: Significant connectivity will be seen between these early regions at times appropriate to the Adolphs model. Particularly, the striate regions will be seen to influence the fusiform gyrus and superior temporal sulcus. As in the Haxby, Hoffman, & Gobbini model, the fusiform gyrus and superior temporal sulcus may demonstrate bidirectional activation.

Chapter 2: Methods

This chapter describes the methods and procedures used in this dissertation. Broadly speaking, the research undertaken in this work consists of three major parts: 1) the construction of a software pipeline to accomplish the proposed non-stationary, group, Granger connectivity analysis, 2) the use of simulations to illustrate connectivity in the time-frequency domain and to validate the newly developed pipeline and, 3) group analysis of MEG data collected from the facial emotion protocol. These three items, as well as their accompanying subtasks, dictate the structure of the Methods, Results, and Discussion chapters which follow.

The work presented here is conceptually simple, but can easily seem complicated when a larger context is not provided. To that end it is important to make clear that the software analysis pipeline is largely assembled from techniques and software which are already established, namely the MxNE family of solvers (Gramfort, Strohmeier, Haueisen, Hamaleinen, & Kowalski, 2011; Gramfort, Kowalski, & Hamalainen, 2012; Gramfort, Strohmeier, Haueisen, Hamalainen, & Kowalski, 2013; Strohmeier, Haueisen, & Gramfort, 2014), and non-parametric Granger causality in the time-frequency domain based upon the Wilson spectral matrix decomposition algorithm (Dhamala, Rangarajan, & Ding, 2008a; Dhamala, Rangarajan, & Ding, 2008b; Wilson, 1972; Wilson, 1978). Two new techniques are developed here but each is also reasonably simple. The irMxNE solver is extended to groups of subjects (Group-irMxNE), but this is done on a concise functional boundary and the correctness is easily seen based on the convex nature of the cost function. It is not necessary to modify the numerical minimization algorithm utilized by the MxNE solvers, and the reader need not understand it. A method of obtaining single-trial output from the irMxNE solver is also developed based on multivariate multiple regression, but electromagnetic principles make clear that this is appropriately considered a linear problem (Gramfort, Strohmeier, Haueisen, Hamalainen, & Kowalski, 2013). Inference is accomplished using a resampling procedure based on shuffling source timecourses within trials. This procedure is drawn from the existing GCCA and MVGC toolkits (Barnett & Seth, 2014b; Seth, 2010). The pipeline is validated using intuitive simulations, and is then applied to the group connectivity analysis of MEG evoked response data collected in the well-known facial emotion protocol (Lysne, 2009).

This research is based on the implementation and application of a software analysis pipeline for use with evoked response MEG data. An *evoked response* is a neural response which is evoked by the application of stimulus, such as the viewing of an emotional face. In order to reduce the signal-to-noise ratio of the data, a stimulus is often presented many times. Each application of the stimulus and subsequent recording of the neural response is called a *trial*. These trials are then averaged and, by doing so, aspects of the data that are evoked by the stimulus, and thus which occur reliably in each trial, are reinforced. Parts of the data which do not correspond to the stimulus, considered noise, tend to be diminished during the averaging process. Most analyses of evoked response experiments are conducted on what is most correctly called the *average evoked response*. Many sources simply call this an *average*, and the MNE software confusingly refers to this as *evoked data*, whereas the individual trials are called *epochs*.

MEG data exists at two levels, *sensor-space* and *source-space*. Sensor-space refers directly to the data that is measured by the MEG sensors, whereas actual activity of neural sources in the brain are referred to as source-space. The goal of neuroimaging research is to identify the activity in the brain that correlates to some task being performed by the subject. Source-space activity is responsible for the activity seen by the sensors. Each sensor receives a contribution from each source, and the overall sensor measurement is a simple summation of the individual source contributions. This propagation of source activities is linear, and is governed by the laws of electromagnetics and by the tissue of the head. This propagation from neural sources to MEG sensors is called a *forward solution*, and the translation from source to sensor-space is implemented by a *forward operator*. A more complex problem is deducing the source activity of the brain from the sensor measurements. This problem is underspecified, and requires that additional assumptions be applied. For example, the MNE L2-norm distributed solution (called a “movie” because it is often viewed as one) imposes the restriction that all activity takes place on the cortical surface and that the least–squared error should be minimized. The solution to this problem creates an *inverse operator*, which translates sensor-space data into the source-space. As opposed to the forward operator which may be directly constructed, the inverse operator usually requires a numerical solution.

Before beginning the following sections, which address the construction and testing of the software pipeline in detail, the reader may wish to visit Figure 9. This figure diagrams the order of operations and the flow of data within the pipeline. If the reader is unfamiliar with wavelet, time-frequency spectra, it may be useful to consider Simulation 1 at this point (Figures 12-15). Simulation 2 provides a simple example of Granger spectral influence between realistic waveforms, and may also be helpful to reference (Figures 16-19). The final group source localizations and Granger spectra for the facial emotion data are found in Figures 36 and 37, and usefully illustrate the end product of this pipeline.

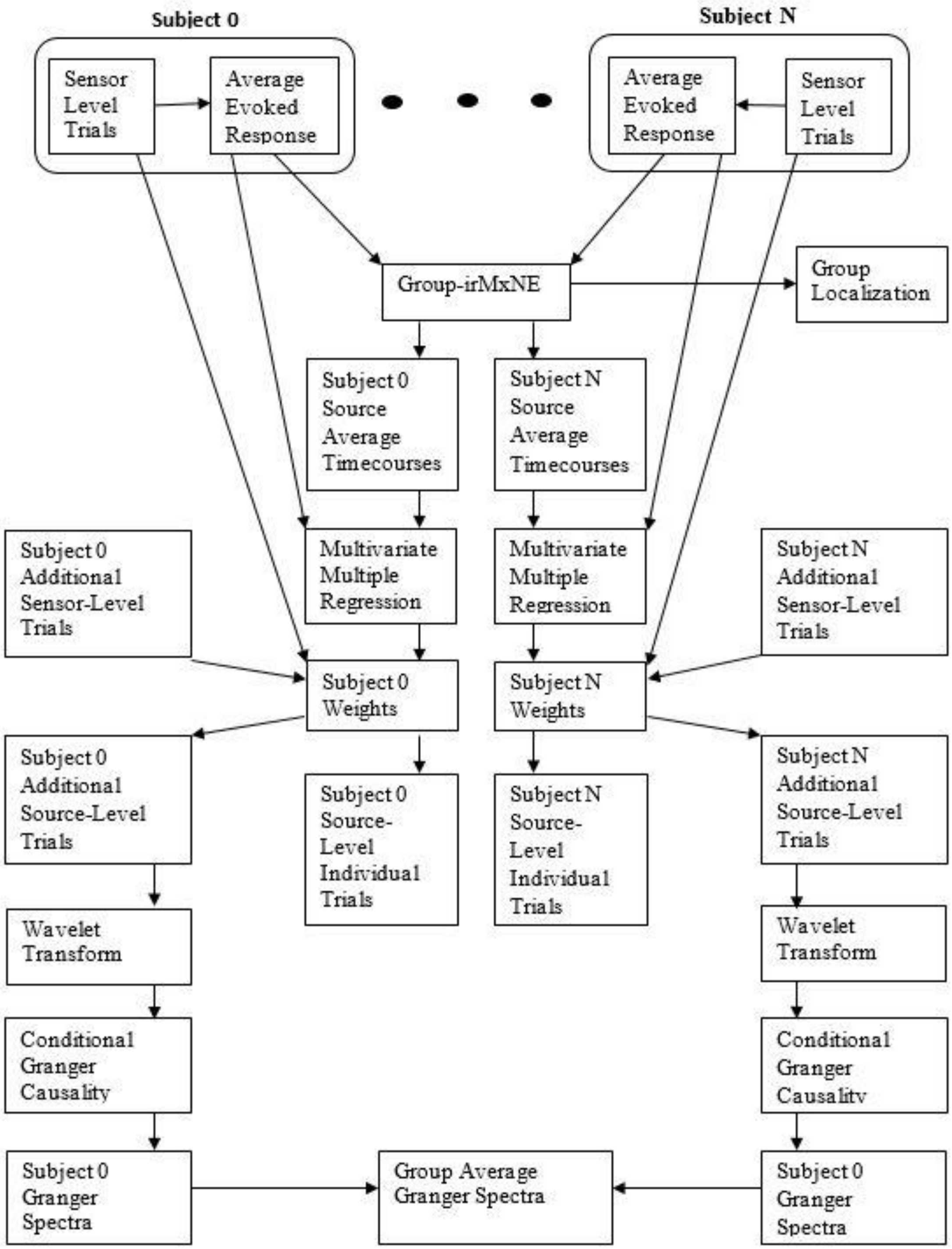


Figure 9: Diagram of the Analysis Pipeline

Multiple subjects' trials and average response are shown entering the pipeline as they are submitted to Group-irMxNE (top). Group-irMxNE generates a group localization and individual source timecourses (necessitated by individual whitening and dewhitening operations). Multivariate multiple regression generates a set of weights for each subject which are then used to propagate sensor-level individual trials to the source-level (both original and any set of additional trials desired). Source-space trials for each subject are then submitted to the wavelet transform and spectral conditional Granger causality estimation. Finally, the Granger spectra are averaged into a true group connectivity spectra (bottom).

Methods for Software

Since this work represents a novel research application, it was unlikely that software currently existed to complete the analysis. Therefore it was determined that software components would be assembled from multiple sources and modified and extended as needed to create the analysis pipeline. To accomplish the goals of this work, a group MEG localization algorithm which produces a limited number of sources would be followed by connectivity estimation and inference. Since neuroimaging analyses are often complex, with numerous options and multiple subjects, it is desirable that individual software components be suitable for use in an automated manner which allows more reliable replication across subjects. The subsections below detail the choices of software and the modifications and extensions to each, and the complete pipeline is diagrammed in Figure 9. To begin with, a broad set of criteria were applied to the software selection process (note that these criteria exclude most commercial software):

1. The software must be freely available (it is possible to include programs for the Matlab environment here because of the UNM student license).
2. The source code and build process must be accessible.
3. The software must have a command line interface, rendering it usable in a scripted environment.
4. It is preferred that the software operate in the Linux environment.

Group Localization

In order to be used in this work, the localization algorithm must produce a model of neural function with the following properties:

1. The generated model must contain a limited number of sources in order for connectivity analysis between these sources to be feasible.
2. The output must include individual trial timecourses, in the source-space, as is required for input to spectral and Granger causality estimation.
3. The output timecourses must not be altered from their original source form, as Granger causality requires the fine resolution of the original, source timecourses.
4. If the localization algorithm does not already support a group model, it must be extensible to one. Group support is desirable at this level, versus later in the pipeline, to avoid the problem of comparing disparate networks between subjects.

In addition, the localization function must be able to operate on data with properties which are typical in neuroimaging:

1. Temporal non-stationarity, characteristic of dynamic neural systems.
2. Temporally dependent timecourses, which are expected in an interconnected system.

averaging over regions.

One of the most rudimentary methods of obtaining source timecourses is to perform a distributed solution followed by simple spatial averaging over all of the vertices within some defined region. This results in a single timecourse representing that region. For an example of this see Lysne (2009). The researcher must trust that the activity of interest lies somewhere within the region, as this is not a true free-form localization. Unfortunately this is not guaranteed, and it is also possible that multiple sources engaged in simultaneous but differing activity are located within the region. Additionally, given the folded nature of the cortical surface, the sources within even a small patch of cortex are likely to have vastly differing magnetic orientations – potentially even canceling one another's contributions. While this technique is useful for gross estimation of the activity within an area, these shortcomings render it unlikely to produce timecourses with fidelity sufficient for follow-on processing.

equivalent current dipoles.

The placement of equivalent current dipoles was briefly considered as a localization option, but no software exists to perform this placement in a group fashion. Theoretically, software could be modified to do so as long as concerns about stationarity were addressed (Gramfort, Strohmeier, Haueisen, Hamalainen, & Kowalski, 2013). Since there was no obvious path forward, this option was rejected. Note, however, that if the analysis involved a preplanned placement of sources versus a free-form localization, then equivalent current dipoles would represent a more attractive option.

components analysis.

A considerable amount of work has been done in the application of components analysis to neuroimaging data, and this appeared to be an early option. In electromagnetic imaging, components analysis is commonly applied in the temporal dimension, whereas in hemodynamic research the spatial dimension is most common. Remembering that components analysis procedures all require the number of sources (or latent variables) to be less than the number of observations, it becomes clear that in our work a temporal application is required. In other words, we anticipate the number of neural sources to be less than the number of temporal samples in the data to be analyzed. This is particularly true of a possible group solution where a small number of sources would be dwarfed by the combined number of observations. At the time, the rejection of ICA for use in this work was a setback, because the only group model known to exist in neuroimaging is the group ICA of Calhoun et al. (Eichele, Rachakonda, Brakedal, Eikeland, & Calhoun, 2011; Calhoun V. D., Adali, Pearlson, & Pekar, 2001).

irMxNE.

The recent development of the *mixed-norm (MxNE)* family of solvers in the MNE-Python package is promising in the current context because these sparse solvers have been specifically designed to work with non-stationary data (Gramfort, et al., 2013; Gramfort, et al., 2014). This family began with MxNE (Gramfort, Kowalski, & Hamalainen, 2012), progressed to *time-frequency mixed-norm estimate (TF-MxNE)* (Gramfort, Strohmeier, Haueisen, Hamaleinen, & Kowalski, 2011; Gramfort, Strohmeier, Haueisen, Hamalainen, & Kowalski, 2013) and, most recently the *iterative reweighted mixed-norm estimate (irMxNE)*, based on the original MxNE (Strohmeier, Haueisen, & Gramfort, 2014; Strohmeier, 2014).

These solvers model the MEG sensors and their relation to neural sources with Equation 5 (where bold, capital letters represent matrices) (Strohmeier, Haueisen, & Gramfort, 2014):

$$\boldsymbol{M} = \boldsymbol{GX} + \boldsymbol{E} \qquad \textit{(Equation 5)}$$

$\boldsymbol{M}$ are the sensor timecourses (averaged over trials), with dimensions NxT, where N is the number of sensors and T is the number of samples (time points, observations). $\boldsymbol{X}$ is the source activity, with dimensions PxT, where P gives the number of sources. The sensor timecourses are related to the sources timecourses by $\boldsymbol{G}$, the gain matrix (mixing matrix, forward operator), with dimensions NxP. Finally $\boldsymbol{E}$, with dimensions NxT, represents additive, white, Gaussian measurement noise: $\boldsymbol{E}\sim\mathrm{N}(0,\lambda\boldsymbol{I})$. (In order to avoid discontinuities at the zero crossings of the output timecourses, irMxNE is necessarily used in this work with fixed orientation sources, implying one source per vertex. Because of this, discussion of the variable *O*, representing the number of dipoles per source location in several of the cited publications, is omitted here.)

Common MEG systems include a few hundred sensors whereas, even when the neural solution space is limited to the cortex, the cortical surface model often contains several thousand vertices. Due to this, Equation 5 is underspecified and there is no single, unique solution $\boldsymbol{X}$. For this reason, additional constrains on the solution space must be imposed. A cost function $F(\boldsymbol{X})$ is constructed and minimized by numerical methods as seen in Equation 6. (The solution to the following equation is $\boldsymbol{X}$^, or the estimated values of source matrix $\boldsymbol{X}$ which minimizes the portion of the equation to the right of the "*arg min"* operator.)

$$\boldsymbol{X}\hat{} = \frac{arg\,min}{\boldsymbol{X}} F(\boldsymbol{X}) = \frac{arg\,min}{\boldsymbol{X}} (f_1(\boldsymbol{X}) + \lambda f_2(\boldsymbol{X})) \qquad \textit{(Equation 6)}$$

In the development of the MxNE solvers, the cost function is comprised of two portions, f_1 and f_2. The first (f_1) is a *data fit term* and the second (f_2), weighted by a *regularization parameter* (λ), is a *regularization*, or *penalty term*.

A norm is a mathematical operator which assigns a positive magnitude to a vector or a matrix. In its most general form, the matrix LP-norm is given by Equation 7 (where *i* and *j* index the rows and columns of the matrix):

$$|| X ||_p = (\sum_{i=1}^{I} \sum_{j=1}^{J} | \boldsymbol{X}_{ij} |^p)^{1/p} \qquad \textit{(Equation 7)}$$

The cost functions of the MNE solvers are constructed on the basis of norms, and the f_1, or data fit portion of the cost function, is always based upon the L2-norm, also referred to as the *Frobenius* norm. The L2-norm distributed solution from which the MNE package originally draws its name (*Minimum Norm Estimation*) minimizes the mean squared error of estimation without the application of a penalty term. This corresponds to Equation 6 without the inclusion of f_2 as rewritten in Equation 8 (where matrix $\boldsymbol{R}$ represents the sensor-space residuals, and $R_{p,t}$ these same residuals in pointwise form).

$$F(\boldsymbol{X}) = f_1(\boldsymbol{X}) = \frac{1}{2} || \boldsymbol{M} - \boldsymbol{GX} ||_{Fro}^{2} = \frac{1}{2} || \boldsymbol{R} ||_2^2 = \frac{1}{2} \sum_{p=1}^{P} \sum_{t=1}^{T} R^2_{p,t} \qquad \textit{(Equation 8)}$$

Note that the squared L2-matrix norm is used here, and the MNE cost function is more easily understood as the simple sum of the squared residuals. The term '$\frac{1}{2}$' is only included for convenience when deriving the optimization method (Gramfort, Kowalski,

& Hamalainen, 2012). The L2-norm solution represents a distributed localization because it is common for nearly every source to have a non-zero level of activation, and the spatial extent of regions of activation is systematically overestimated (Gramfort, Kowalski, & Hamalainen, 2012). Figure 4 shows the L2-norm distributed solution for a single subject at a single time point while viewing an emotional face, and frames from the facial emotion average L2-norm solution are seen in Appendix D.

The MxNE solvers include a penalty term in their cost functions, and this term is represented by a *mixed norm*. Norms of differing order offer different properties, and the purpose of a mixed norm is to foster different properties along separate dimensions of a matrix. As opposed to the data fit term, which is applied to the sensor-space residuals, the mixed norm is applied to the source activity estimate. In particular, the MxNE solver utilizes the L21-mixed-norm seen in Equation 9.

$$f_2(\boldsymbol{X}) = \sum_{p=1}^{P} \sqrt{\sum_{t=1}^{T} \boldsymbol{X}_{p,t}^{2}} \qquad \textit{(Equation 9)}$$

The L21-mixed-norm applies the L2-norm behavior to the time dimension of the source activity, and L1-norm to the spatial dimension. Whereas the L2-norm tends to promote a smooth, distributed solution as discussed above, the L1-norm promotes sparsity, with a relatively small number of non-zero entries in the resulting solution. As applied to the source-space by MxNE, this results in a limited number of active sources (with all others being set to zero for the duration of the response) and, for those sources which are active, smooth and continuous timecourses throughout the response. A graphical example of this behavior may be seen in Fig. 1 of Gramfort, Strohmeier, Haueisen, and Kowalski (2011).

The MxNE solvers allow prior weights to be applied to the source matrix before minimization of the cost function as seen in Equation 10. These weights are commonly derived from fMRI data, or from the L2-norm solution (as is used in this work).

$$f_2(\boldsymbol{X}) = \sum_{p=1}^{P} \sqrt{w \sum_{t=1}^{T} \boldsymbol{X}_{p,t}^{2}} \qquad \textit{(Equation 10)}$$

The MxNE solver utilizes a numerical algorithm to minimize the cost function constructed from a sum of f_1 and f_2 (Equation 6). This algorithm necessitates a requirement that both f_1 and f_2 be convex functions and, given this, a globally optimal solution is ensured regardless of the initialization state of the solver. In practicality, this means that the MxNE solver deterministically returns an optimal solution, and that it need not be run multiple times.

Unfortunately, the MxNE solver is noted to produce non-optimal source localizations and biased amplitude estimates and, for these reasons, is superseded by the irMxNE algorithm (Strohmeier, 2014; Strohmeier, Haueisen, & Gramfort, 2014). The irMxNE algorithm replaces the L21-mixed-norm in the f_2 portion of the cost function with a block application of the L2,0.5-quasi-norm. This application depends upon multiple sources being present at each vertex, and $\boldsymbol{X}_p$ is used to denote the block of sources associated with vertex p. However, when the number of sources per location is set to '1', it may be seen that Equation 11 is very similar to the previously seen Equation 9 (with the difference being an extra square-root operation).

$$f_2(\boldsymbol{X}) = \sum_{p=1}^{P} \sqrt{\left|\left|\boldsymbol{X_p}\right|\right|_{Fro}} \qquad \textit{(Equation 11)}$$

Despite being similar in the present case, the f_2 term of the irMxNE cost function may no longer be assumed to be convex, and a new numerical minimization algorithm is employed. This algorithm introduces an additional level of looping and, in order to assure convergence, the first iteration is performed by the original MxNE procedure (i.e. the initialization point of the irMxNE solver is the MxNE solution for the given data). The solver then iterates as follows in order to minimize Equation 12 (with iterations indexed by variable k):

$$\boldsymbol{X}^{\wedge(\boldsymbol{k})} = \frac{arg\,min}{\boldsymbol{X}}\ \frac{1}{2}\,||\ \boldsymbol{M} - \boldsymbol{GX}\ ||_{Fro}^{2} + \lambda\frac{1}{2}\sum_{p=1}^{P} \frac{||\boldsymbol{X_s}||_{Fro}}{\sqrt{\left|\left|\boldsymbol{X_S}^{\wedge(k-1)}\right|\right|_{Fro}}} \qquad \textit{(Equation 12)}$$

Where $\boldsymbol{X}_s$ denotes the row of $\boldsymbol{X}$ corresponding to source s, and $\boldsymbol{X}^{\wedge(\boldsymbol{k})}$ refers to the estimated solution $\boldsymbol{X}^{\wedge}$ at iteration k. (Note that in Equation 12 '*S*' is used to denote sources, which is done here to maintain consistency with the original authors' form of these equations (Strohmeier, Haueisen, & Gramfort, 2014). Hence, in the next section where these methods are extended to multiple subjects, a different subscript than *S* is used to denote subjects.)

Group-irMxNE.

With regard to the group solution, a question commonly arises as to why the sensor-level data from multiple subjects cannot simply be averaged or concatenated prior to submission to the solver. The rejection of this approach lies in the different head and brain shapes of the subjects as well as the subjects' differing head positions within the MEG sensor array. The solution to these problems is to morph the individual head topology into a common source-space and then to use an individual forward operator for each in order to correct for the differing head positions. (As opposed to MEG, EEG data is collected directly from the scalp using a standard electrode placement. This solves both the head shape and position problems and averaging or concatenating is appropriate.)

Whereas the morphing procedure is done independently of localization, the unique forward operators are a direct input to this group localization. A possible solution to the problem of differing forward operators would be to employ an algorithm which could translate the head position of each subject to a common location. This involves a complex electromagnetic adjustment however, and a comprehensive search revealed that no suitable program was available for use in this work.

None of the MxNE family of solvers provides a group solution. However, their cost function implies that one is possible. The solvers minimize the previously seen Equation (6), which is restated in Equation (13). The difference between the MxNE and irMxNE cost functions lies in the penalty term given by f_2. In MxNE both f_1 and f_2 are required to be convex, but in irMxNE, f_2 no longer is so.

$$\boldsymbol{X}^\wedge = \frac{arg\ min}{\boldsymbol{X}} f_1(\boldsymbol{X}) + \lambda f_2(\boldsymbol{X}) = \frac{arg\ min}{\boldsymbol{X}} \frac{1}{2} \left|\left| \boldsymbol{M} - \boldsymbol{GX} \right|\right|_{Fro}^2 + \lambda f_2(\boldsymbol{X}) \quad \textit{(Equation 13)}$$

A group solution suggests itself in the form of Equation 14. In this case the $f_1(\boldsymbol{X})$ term seen in Equation 12 has changed from the sum of squared errors for a single subject, to the combined sum of squared errors for a group or cohort consisting of multiple (C) subjects. Since the sum of squared errors is a convex function for each subject individually, their combined, positive sum is also convex, and the new function may be submitted to the MxNE minimization procedure as before. The f_2 mixed- and quasi-norm term, calculated on source estimate $\boldsymbol{X}$, which is now the source estimated activity for the group model, does not change. Given that the group f_1 term remains suitable for the MxNE minimization algorithm, it is reasoned that it continues to be suitable for the irMxNE procedure as well. The group cost function represented by Equation 14, with the irMxNE quasi-norm, may then be submitted to the irMxNE solver as before.

$$\boldsymbol{X}\hat{} = \underset{\boldsymbol{X}}{\arg\min} \sum_{c=1}^{C} \left\{ \frac{1}{2C} \, || \, \boldsymbol{M} - \boldsymbol{G}\boldsymbol{X} \, ||_{Fro}^{2} \right\} + \lambda \, f_2(\boldsymbol{X}) \qquad \textit{(Equation 14)}$$

Advantageously, this new cost function represents a concise functional boundary within the solver which may be separated cleanly from the minimization procedure and from the mixed- and quasi-norms (f_2). The single-subject version of the solver iterates back-and-forth between the sensor and source-spaces, with each iteration refining the source estimate, using the gain matrix to translate the new estimate to sensor-space, and calculating the new value of the cost function. The new cost value is used to inform the next iteration, and the process continues until either a minimum tolerance threshold is met or until the maximum number of iterations is reached. Fortuitously, adapting this algorithm to multiple subjects is a simple matter of applying the source estimate to multiple subjects and averaging their cost values at each iteration.

Equation 1 stipulates that the measurement noise ($\boldsymbol{E}$) must be uncorrelated. The MxNE solvers accomplish this by applying a whitening procedure to the data before the solver begins to operate on it, and a dewhitening step before returning the results to the user. These operations are unique to the individual subjects (Alexandre Gramfort, personal communications, October, 2014). For this reason, the results returned from Group-irMxNE are comprised of a single set of localizations, and a set of source timecourses for each subject. For examples of Group-irMxNE output, see Simulations 5 and 6.

The original Python code of the MxNE solver is available within the software distribution, and it was undertaken to modify this code for the group solver. The following functions were modified (with their file name and path relative to the top of the MNE-Python source tree given in parenthesis):

• *dgap_l21*	*(mne.inverse_sparse.mxne_optim)*
• *mixed_norm_solver_prox*	*(mne.inverse_sparse.mxne_optim)*
• *mixed_norm_solver*	*(mne.inverse_sparse.mxne_optim)*
• *mixed_norm*	*(mne.inverse_sparse.mxne_inverse)*

As discussed previously in this section, the application of Group-irMxNE requires individual subjects to be morphed into a common source-space prior to being invoked (the group solution is, by definition, in a single, common, source-space). In multi-subject neuroimaging studies, subjects' brain topology is often morphed onto that of either an average brain, a single, chosen subject of the study, or an average of the subjects involved. Unfortunately the subject-morphing algorithm which is advertised to exist in the MNE software (program *mne_setup_source_space* option *–morph*) was found to be inoperable. Attempting to morph subjects in the facial emotion dataset used here onto each other or onto an average brain resulted in an error of multiple source vertices being mapped to a single target vertex. Consultation with the MNE developers as well as investigation of the source code for this program revealed repairing this problem to be beyond the scope of this work. (MNE does support this functionality in the *mne_make_movie* program as well but, in this case the individual subjects' distributed localizations are morphed after they are complete – an option which occurs too late in the analysis pipeline to be part of a group solution.) It remains possible to utilize Group-irMxNE, but requires that the MEG data of the subjects be registered directly onto an average brain without the benefit of a morphing option.

Group-irMxNE testing.

Primary testing of Group-irMxNE is performed by submitting individual subjects to the group program and comparing their results to that of the original, single-subject irMxNE code. Additionally, it is reasoned that the optimal solution for a single subject should be the same as the solution for multiple instances of that same subject. (This degenerate case tests the averaging procedure of the sensor-space squared errors over multiple subjects.) Several instances of the same subject were submitted to Group-irMxNE at once and compared to the output for these subjects from the single-subject code. At the suggestion of Alexandre Gramfort and Dennis Strohmeier (personal communications, April, 2015), additional testing was performed at the level of the complete pipeline (Group-irMxNE plus the single trials multivariate regression and Granger causality estimation), and done with simulated data (see methods, results, and discussion for Simulations 3-6).

irMxNE single trials regression.

In addition to the lack of a group solution, the MxNE solvers do not produce single-trial output. All implementations accept an average sensor-level neural response as input (evoked data, in MNE terms), and produce source-level timecourses for the localized sources corresponding to the average input. In other words, the sensor-level trials must be averaged prior to submission and the output timecourses correspond to the average input, not to the original individual trials.

Fortunately a solution to this problem is suggested in the introduction of Gramfort et al. (2013), where it is explained that the electromagnetic relationship between sources in the brain and the measurements at the sensors is necessarily a linear one. In MEG, each neural source is visible to all of the sensors, and the measurements of the sensors represent a simple summation of the combined source activity. Additionally, changes in the sensor measurements are related in a linear fashion to the activity of the sources. Such a linear system is a candidate to be solved with multivariate multiple regression. The sensor measurements represent the predictors, the source values the criteria, and the observations are collected over time. The regression is implemented in Python, using the Statsmodel package of Seabold and Perktold (2010).

irMxNE single trials regression testing.

The single trials regression is tested by translating the individual trials back to sensor-space using the forward operator, averaging them, and computing the resulting model's goodness of fit to the similarly processed irMxNE average timecourses (for an example, see Table 1). Additionally, the source-space average of these trials is plotted along with the irMxNE timecourses for visual inspection (Figures 20 b), 23 b), 26 a), c), 30 a) c) e)).

Connectivity

It was anticipated in advance that Granger causality would be used for the connectivity portion of the pipeline unless it proved to be unsuitable. This decision was made based on its maturity, flexibility, and utility across a wide range of disciplines. However, the needs of this work impose a pair of related requirements which are not commonly associated with Granger causality:

1. The connectivity algorithm must operate on non-stationary systems.
2. Since non-stationary analysis precludes the use of the ubiquitous Fourier frequency domain, the connectivity algorithm must be suitable for use with wavelet spectra in the time-frequency domain.

Many neuroimaging researchers believe that Granger causality is limited to MVAR-based analysis of stationary systems. The origin of this belief may lie in

Granger's original development, which is based exclusively on MVAR concepts and their analogs in the Fourier domain (1969). Fortunately, recent publications by Dhamala, Rangarajan, and Ding demonstrate the applicability of non-parametric Granger causality to non-stationary neuroimaging data in the time-frequency domain (2008a; 2008b).

Fieldtrip conditional Granger causality.

Although the principles of Granger causality are well understood, its implementation, particularly of conditional Granger causality in the frequency domain, is complex (Ding, Chen, & Bressler, 2006). This is true of the numerical Wilson spectral matrix decomposition algorithm as well (Wilson, 1972; Wilson, 1978). Fortunately for the progression of this work, an implementation of non-parametric conditional Granger causality based on the Wilson decomposition is found in the Fieldtrip software for the Matlab environment (Oostenveld, Fries, Maris, & Schoffelen, 2011). The software is incomplete however, and the interface is largely undocumented. File and function headers indicate that the implementation of the Wilson algorithm is provided by Dhamala and Rangarajan (functions *sfactorization_wilson* and *sfactorization_wilson2x2* in Matlab source files of the same names). Authorship of the non-parametric Granger causality functions, which rely upon the Wilson algorithm, is less clear. Although the code is very cleanly written, ownership, which is usually claimed in the file headers, is missing. Questions about this code remain unanswered on the Fieldtrip e-mail forum as well. Nonetheless, it was discovered through examination of the code that this functionality could be accessed via the *ft_connectivity_analysis* interface, and testing reveals that it works well.

Granger causality extensions to the wavelet domain.

Although the analysis of non-stationary systems, as is enabled by non-parametric Granger causality, depends upon the two-dimensional spectra of the time-frequency domain, the *ft_connectivity_analysis* interface only allows the submission of one-dimensional spectra (i.e. Fourier spectra). Since the works of Dhamala, Rangarajan, and Ding show that the combined Wilson decomposition and non-parametric Granger causality may be applied to the wavelet domain as well, this is judged to be an oversight in the Fieldtrip software. This shortcoming is easily remedied by calling *ft_connectivity_analysis* once for each time sample in the spectra.

A Matlab wrapper function is developed for *ft_connectivity_analysis* which takes as input the individual-trial source timecourses, generates the wavelet power and cross power spectra for each trial, and averages these spectra. This follows the spectral estimation procedure of Dhamala, Rangarajan, and Ding. (A question commonly arises as to whether the single-trial spectra must be averaged, or if the single trials themselves may be averaged in the time domain and the wavelet transform taken once upon the resulting average timecourses. This would be particularly advantageous since irMxNE already outputs average timecourses, and thus the multivariate regression procedure to produce single trials could be avoided. Unfortunately, although the wavelet transform is itself linear, the operation of multiplying each spectral value by its complex conjugate to produce power and cross power spectra is not. The results are not equivalent and, when created this way, the spectra cause the Wilson decomposition algorithm to fail with a matrix inversion error.)

The average spectra created this way are then submitted, time point by time point, to Granger estimation using *ft_connectivity_analysis*. Since the processing at each time point is independent of all others, this problem submits itself easily to parallel computation, which is accomplished using the *parfor* looping construct in the Matlab parallel toolkit. This is seen to decrease running time dramatically based on the hardware available, which becomes especially important when the resampling-based inference procedure is employed. An additional level of loop exists in this wrapper to implement the multiple reshuffling of trials needed by the resampling inference procedure discussed below.

Although wavelet transforms are available in a variety of packages, a custom implementation of the complex Morlet wavelet (Torrence & Compo, 1998) was used here in order to apply the scaling of Liu et al. (2007) and Veleda et al. (2012)(Figure 5). This scaling causes the peak power values in the wavelet spectra to remain constant across frequencies, assisting with interpretation (Figure 6). The wavelet parameter $w_0=12$ was used for improved spectral resolution, although coming at the expense of temporal resolution (Mukesh Dhamala, personal communications, October, 2013). Unlike the Fourier transform, the frequencies present in the wavelet spectra are at the choice of the user. All wavelet spectra in this work are comprised of the integer frequencies between zero and 60 Hz.

Granger causality testing.

The conditional Granger causality program was tested on the bivariate autoregressive system of Dhamala, Rangarajan, and Ding, and compared to these authors' results (2008a; 2008b). This system switches direction of influence at its midpoint, rendering it inherently non-stationary. Results of this testing may be seen in Simulations 3-6.

First half of trial:

$$x_0(t) = 0.55x_0(t-1) - 0.80x_0(t-2) + 0.25x_1(t-1) \tag{Equation 15(a)}$$

$$x_1(t) = 0.55x_1(t-1) - 0.80x_1(t-2) \tag{Equation 16(a)}$$

Second half of trial:

$$x_0(t) = 0.55x_0(t-1) - 0.80x_0(t-2) \tag{Equation 15(b)}$$

$$x_1(t) = 0.55x_1(t-1) - 0.80x_1(t-2) + 0.25x_0(t-1) \tag{Equation 16(b)}$$

Goodness of Fit

irMxNE.

Goodness-of-fit of an irMxNE model may be calculated using a procedure suggested by Gramfort (personal communication, October 3, 2014, via the MNE e-mail list). The input evoked data is serialized into a single vector of length *(#sensors × #samples)*. The average source timecourses output by irMxNE are translated to sensor-space via the forward operator and then serialized the same way. The Pearson correlation is then calculated between the two vectors. This correlation gives a measure of how well the irMxNE model fits the original simulation, and the correlation is then squared, resulting in the percentage of the original variance captured.

Goodness-of-fit of the source-level single trials may be calculated in the same way, with the individual trials being translated to sensor-space, averaged, serialized, and correlated to the original evoked data in the manner above.

Inference

individual Granger causality.

Since the null distribution of the wavelet power spectra can only be estimated (Torrence & Compo, 1998), and is unknown for the measures of Granger influence (Seth, 2010), inference is performed on each of these spectra using a resampling procedure. The timecourses of individual trials of x_0 and x_1 are reordered such that they no longer represent pairs generated together. For example, the x_0 timecourse from trial #1 might be randomly paired with the x_1 timecourse from trial #2. At each reshuffling of the trials the wavelet, Granger, and condition Granger spectra are recalculated. A number of these reshufflings are used to generate a null distribution for each spectra at each time point and frequency (i.e. a complete null spectra). This resampling procedure is adapted from the GCCA Toolkit of Seth (2010). The algorithm is modified here such that a randomly occurring correct pairing of the x_0 and x_1 trials, however unlikely, is prohibited from taking place. The percentile of the correctly paired spectral data against these null distributions is then used to determine the *p* value at each spectral position.

Resampling procedures are commonly employed on a pointwise basis, and the trialwise algorithm employed here may seem confusing. The benefit of this approach is that the power spectra of the individual sources remain unaffected, and only the cross-power is changed by each reordering. When the decomposed cross-spectra (following the Wilson algorithm) are normalized by the product of the power spectra to form the Granger measure of influence, only the numerator of this ratio changes, with the denominator remaining the same as with the correctly matched timecourses. If pointwise reordering were employed, the power spectra and the cross-power spectra would both be disrupted. The resulting ratio would then be the quotient of two noise processes, with unpredictable results.

A criticism is possible that, the reordered timecourses representing an evoked response, some similarity exists between trials. Even mismatched timecourses may be expected to show some level of connectivity. However, this connectivity should be lesser than that between the correctly matched trials and, at worst, this causes the resampling test proposed here to become more conservative. (This may even be desirable given that computational resources preclude the number of resampling trials needed to accommodate correction for multiple testing.)

For this work, a two-tailed test is employed with α=0.01. For display purposes, each *p*-value is subtracted from '1' and this difference is transformed into a negative value if it falls within the lower tail of the null distribution. A significance spectra is then created such that non-significant *p*-values are set to zero. When displayed as an image, the grey background color indicates locations in the spectra which do not significantly differ from the empirical null distribution. Those locations where the spectra is significantly greater than the null (at α/2 since this is a two-tailed test) are shown in red, and those significantly below the null in blue. 250 resamplings are performed for each spectra.

group Granger causality.

It was originally intended that a method for group inference would be developed for use on the Granger causality estimates following Group-irMxNE. Under this method, each subject contributes their individual Granger output spectra, and these spectra would be used as observations in an inference procedure (for the individual spectra output from Group-irMxNE applied to the facial emotion data, see Appendix C). Inference could be performed on a point-wise, frequency-wise basis across the collection of individual spectra to produce a *p*-value at each time and frequency. Several issues would need to be addressed. First, although frequency-domain Granger connectivity values are known to be positive, their distribution is unknown (Seth, 2010). Even without knowing this distribution, and particularly its analytic form, we might assume that parameters of some type partially determine its shape. A theoretical question arises as to how these parameters could be normalized across subjects. In other words, how could we establish that all subjects' Granger spectral values were actually members of the same distribution with the same parameters? Furthermore, the estimation of such parametric values is often the origin of the stationarity requirement (for example, the mean and variance of a time series being estimated by collecting samples over time). Such limitations would need to be avoided. Resampling algorithms as described above are one possibility.

Granger causality spectra are the principal focus of this research, and less emphasis is placed on wavelet power and cross power spectra (because they are unrelated to significance of the connectivity results).

Unfortunately, time constraints did not allow group inference to be addressed in this work.

Methods for Simulations

This work presents a set of six simulations with the twofold goals of introducing the reader to time-frequency spectra and its interpretation, and validating the software pipeline described above. The simulations are as follows:

1. *Simulation 1* introduces the wavelet power and cross power spectra and contrasts them with the better-known Fourier spectra. Results for a simple, known system are presented in both domains for comparison.
2. *Simulation 2* extends the application of spectral decomposition and Granger causality to realistic MEG-derived neuroimaging signals. This simulation validates techniques used here on waveforms derived from actual experimental data, as opposed to data generated from simple autoregressive systems (as are used in the remaining simulations). These results also identify the pattern of Granger spectral connectivity to be expected when influence between sources is encountered in the facial emotion data.
3. *Simulation 3* tests the complete pipeline end-to-end with simulated data for a single subject. Two cortical sources are activated by the direction-of-influence-switching, nonstationary, dynamic MVAR system of Dhamala, Rangarajan, and Ding (2008a; 2008b).
4. *Simulation 4* tests the pipeline end-to-end using a single subject with more complex activity comprised of two independent pairs of sources similar to Simulation 3. The pairs of sources are located in such a way as to be spatially distinct from one another.
5. *Simulation 5* tests the pipeline for two subjects, each containing one of the pairs of sources from Simulation 4. This simulation not only tests the pipeline for use with multiple subjects but, since the sources presented are spatially distinct, the ability of the pipeline to preserve intuitively separate activity is also tested.
6. Simulation 6 tests the pipeline using three subjects, each with a similar pattern of neural activity. The ability of the pipeline to generalize these similar subjects into a group model is assessed.

Methods for Simulation 1

In this simulation a pair of 300-sample timecourses are created, where each contains pure sinusoidal components at 20, 40, 60, and 80 Hz. In the first timecourse, x_1, the 20 and 60 Hz sinusoids have amplitude ‘0.5’ and the 40 and 80 Hz have amplitude ‘1’. In the second, x_2, the 40 Hz sinusoid has amplitude ‘1’ with all others being ‘0.5’. The phase of all sinusoids is zero and the sampling frequency is taken to be 200 Hz.

The discrete Fourier and Morlet wavelet (w_0=12) transforms are taken of this pair of signals and the power and cross power spectra are created. The power spectra of x_1 and x_2 is simply the pointwise squared magnitude of each (alternatively, the pointwise product of each with its own complex conjugate), and the cross power spectra is the product of x_1 with the conjugate of x_2 and vice versa. The wavelet power spectra are then averaged over time in order to create plots which are more directly comparable to the Fourier power spectra.

Results for Simulation 1 are seen in Figures 12-15.

Methods for Simulation 2

This simulation extends the application of Granger causality from simple, purely MVAR systems to more realistic MEG timecourses. The purpose of this is to identify the spectral pattern of connectivity expected to occur in actual MEG data when influence between sources is present. A pair of waveforms are selected from an analysis of median nerve data (although any shapes would work). One of these shapes is designated x_0 and the other x_1, and a system is constructed such that x_0 appears to cause x_1. 500 trials are constructed from these sample waveforms by scaling the waves in time and amplitude. For each trial a group of four scaling factors are generated from a normal distribution, and all four values are correlated at $r=0.8$. These values are used to scale the length and amplitude of x_0 and x_1, and x_1 is set to begin at the completion of x_0. The end result of this is trials in which the amplitude and length of x_0 are predictive of the onset, amplitude, and length of x_1. These trials are then submitted to the Granger estimation program constructed above. Results for Simulation 2 may be found in Figures 16-19.

Methods for Simulation 3

This simulation tests the entire data processing pipeline of this work for a single subject from end-to-end, and is diagrammed in Figure 10. The methods and results of this simulation are described in detail in the paragraphs below as they apply directly to Simulations 4-6 as well as to understanding the processing of the facial emotion data.

Three-hundred paired trials of two time series, x_0 and x_1, are generated according to Equations 15 and 16. Each trial is two-hundred samples in length, and sampling is asserted to occur at 200 Hz. This represents one second of time, or zero to 1,000 milliseconds for each trial. The noise variance of both series is set to '1e-9' and the covariance to '0' (noise in this context refers to the random innovations which energize the system, and the scaling is selected such that the resulting signal amplitudes are realistic for MEG data). The direction of influence between the trials is initially $x_0 \rightarrow x_1$, and this direction is reversed at the midpoint of each trial. The trials are created using the Nitime package for performing time series operations in neuroimaging analysis (Rokem & Perez, 2009). The individual simulated trials as well as their average may be seen in panel c) of Figure 20. (Figure 10, Step 0)

The MNE-Python simulation utility *generate_sparse_stc* is used to choose a pair of source vertices on the surface of a test cortex, with one source randomly selected in each of the left and right auditory areas (Figure 20, panel a). The test subject utilized here is the 'sample' subject from the dataset provided with the MNE software distribution. A forward operator is created for this subject using '*oct-3'* spatial sampling, which results in the left and right cortical surfaces being represented by a mesh containing a relatively small number of 1,284 vertices apiece. (By using a limited number of cortical vertices it is hoped to make each potential source more distinct from its neighbors as well as to reduce the running time of the solver.) The potential equivalent current dipole sources at the cortical vertices are placed into *fixed* orientation, implying that that their orientation remains fixed in a normal position to the surrounding cortical surface at all times. This also means that only a single potential source is placed at each vertex. The covariance matrix from the MNE 'audviz' simulation is adopted for use in this simulation as well.

The paired timecourses for each simulated source-space trial are assigned to the cortical sources created in the previous step (x_0: left, x_1: right). Each pair are then propagated to sensor-space using *generate_evoked*, and noise with *snr=3* is added at the same time. The forward operator and covariance matrix discussed above are used in this operation as well. The sensor-space trials with noise are then averaged to create an evoked response (Figure 10, step 2).

The evoked data as well as the forward operator and covariance matrix are then used to generate the dSPM L2-norm localization for this subject using the utility *mne_make_movie*. Inputs *spm* and *picknormalcomp* are set, as well as *smooth=7* (this step is not shown in Figure 10). This type of distributed localization is commonly referred to as an MNE ‘movie’ and, although it is not shown in this work, it is analogous to the snapshots from the facial emotion group movie seen in Appendix D. This movie is generated here in order to be used as weights by the Group-irMxNE localizer below.

The evoked data, forward operator, covariance matrix, and L2-localization, are input to Group-irMxNE (even though this is a single-subject simulation, the group version of the code is utilized for testing purposes). The localizer is run for 150 iterations and 15 sub-iterations (parameters *maxit* and *n_mxne_iter*). Since the number of sources in the data is known to be two, the localizer is run repeatedly and the regularization parameter (λ) adjusted to a value of 0.05, such that the output solution consists of two sources. As required, the *loose* parameter is set to zero. The *depth* parameter is set to 0.9 as is done in the example of Gramfort et al (2014). Other Group-irMxNE input parameters are set as follows: *tolerance=1e-14*, *pca=False*, *time_pca=False*, *debias=True*, *weights_min=0.01*, and *active_set_size=100*. (Figure 10, step 3)

The multivariate regression procedure is then applied to the input evoked data and the average timecourses output from Group-irMxNE for each of the output sources. In this case, the two average source timecourses, also 200 samples apiece, are regressed onto the 305 sensor timecourses to create two sets of 305 regression weights. These weights may then be applied pointwise to any data in order to ‘propagate’ that data through the Group-irMxNE model. In our case the sensor data for the original 300 trials are propagated to create the corresponding 300 source-level timecourses. This may be thought of as an inverse operator that is specialized to a specific Group-irMxNE model. (Figure 10, step 4)

The original Nitime timecourses and the regression propagated timecourses are input to the Granger connectivity estimation program in Matlab. Input settings include:

w_0*=12*, *MaxFreq=60,* and *FreqStep=1*. The wavelet power, Granger connectivity, and conditional Granger connectivity spectra are seen in the left hand columns of Figures 21 and 22 respectively.

Since the null distribution of the wavelet power spectra can only be estimated (Torrence & Compo, 1998), and is unknown for the measures of Granger influence (Seth, 2010), inference is performed on each of these spectra using the resampling procedure described above. 250 reorderings are performed for each spectra, and a two-tailed test is performed at α=0.01. The results are seen in the right hand columns of Figures 21 and 22. Those locations where the spectra is significantly greater than the null are shown in red, and those significantly below the null in blue.

Goodness-of-fit of the Group-irMxNE and propagated trials models are calculated using the procedures described above. The correlations and squared correlations are seen in Table 1 (‘irMxNE Total’ and ‘Trials Total’ respectively). In addition to the models containing both of the two original sources, goodness-of-fit is calculated for each source alone. These results are also seen in Table 1 (‘Src 0’ and ‘Src 1’ for both the Group-irMxNE and Trials models).

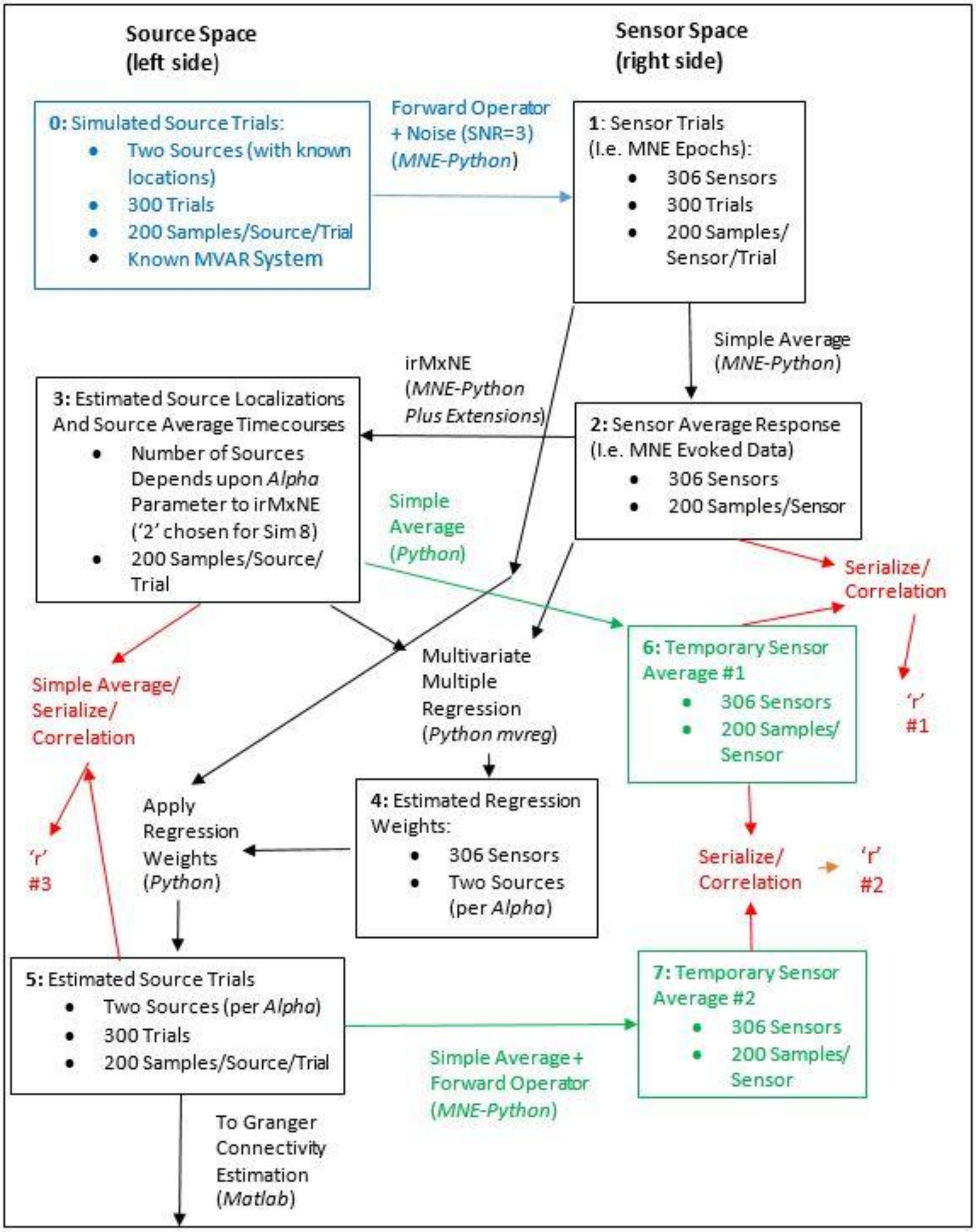


Figure 10: Simulation 3 Data Pipeline

Data generation, application of the data pipeline, and estimation of goodness-of-fit for Simulation 3. The core of the data pipeline is shown in black. Steps employed to simulate MEG data in lieu of an actual subject are shown in blue. Additional steps performed to facilitate goodness-of-fit estimation are shown in red and green. Source-space data is shown on the left side of the figure, and sensor-space on the right.

Methods for Simulation 4

This simulation tests the ability of the pipeline to work with more complex patterns of sources within a single subject. Two pairs of sources are created according to the procedure in Simulation 3, and one pair is assigned to the left and right auditory cortices and the second to the primary visual. The result is a single subject with two pairs of independently interacting sources. The auditory and visual areas are chosen to represent regions distinct from one another which the localizer is hoped to be able to identify without attempting to combine them. Results of Simulation 4 are seen in Figures 23-25 and Table 2.

Methods for Simulation 5

This simulation tests the ability of the pipeline to work with more than one subject with distinct sources. Locations and timecourses from the two pairs of sources in Simulation 4 are reused but, in this case, each pair of sources are assigned to a separate subject. The product of this are two subjects, one with a pair of left and right interacting auditory sources, and the second with a pair of interacting sources in the primary visual cortex. The goal of this simulation is to begin testing Group-irMxNE and the rest of the pipeline on multiple subjects, but with a pattern of sources which may be compared to Simulation 4. Again, the sources in each subject are intended to be distinct from one another, and the group solution should not attempt to combine them. Note that group output from the pipeline consist of a single set of localizations (in this case matching Figure 23, a)), individual subject timecourses (Figure 26), spectra (Figures 27 and 28 for the two subjects respectively), and group average spectra (Figure 29). Goodness-of-fit is similarly decomposed by subject in Table 3.

Methods for Simulation 6

This simulation tests the ability of Group-irMxNE to operate on multiple subjects with collocated sources. Three separate subjects are created as described in Simulation 3. The pair of left and right auditory sources for each subject are randomly selected from the auditory regions, but the sources of the subjects were required to be non-overlapping. This produces three subjects with interacting auditory sources within each subject, but the pairs of sources are independent between subjects.

The auditory regions are selected for all three subjects in order to create subjects with closely located pairs of sources, which the localizer should begin to generalize. As opposed to Simulation 5, where intuition suggests that the activity on the two subjects' cortices is unrelated and should remain distinct, the activity of the three subjects here can be construed as the same cortical process and therefore the localizer should make some attempt to combine it. It is anticipated that the results will be recognizably similar to those seen in Simulation 3 (times three), but that the activity in the sources of the three subjects should begin to merge. The combined localization of all three subjects is seen in Figure 30, their timecourses in Figure 31, the corresponding spectra for each in Figures 32, 33, and 34, and the group average spectra in Figure 35. Goodness of fit information is found in Table 4.

Methods for Facial Emotion Analysis

MEG data from twenty-three subjects collected during a facial emotion protocol is reused here (Lysne, 2009). The data is reprocessed for use in this work, and then submitted to individual and group analysis with the pipeline constructed above. Despite the desired end result being a group analysis, this work recommends that individual analyses be conducted as part of this process for purposes of data cleaning prior to group submission. Additionally, since questions arise in this work regarding the direct co-registration of subjects to the average brain, individual analyses are conducted both using the average brain and individual head models where available (Table 5).

Subjects

Twenty-three subjects were recruited by convenience from the students and staff at the Mind Institute (Albuquerque, NM). Ten subjects were male and thirteen female, with an average age of 48 years. All procedures were approved by the University of New Mexico North Campus Human Research Review Committee (Thoma, 07-179).

Data Collection

Continuous MEG data was collected on a 275-channel CTF system (VSM Medtech, Toronto, BC). Stimulus was comprised of faces depicting seven emotional conditions (happiness, sadness, anger, fear, surprise, disgust, and neutral) drawn from

Ekman and Friesen (1976). Two control conditions, swirled faces and non-human sculptures, were employed as well. Each face was presented for 1.5 seconds, with an interstimulus interval varying randomly between 1 and 3 seconds. Each condition was presented 120 times, for a total of 1080 trials per subject. Data was collected at 600 Hz. Examples stimuli are shown in Figure 11. The stimuli were presented using a projector with an approximate 30ms delay and unknown jitter properties. This delay is uncorrected in the present analysis.

Along with MEG data, a three-dimensional model of the external surface of the head was generated using a Polhemus Fastrak digitizer (Colchester, VT). (These head shape points were collected to facilitate later co-registration between the MEG dataset and the physical, MRI-based head model.)

Several subjects had pre-existing structural MRI scans, and scans were arranged for others. At the time of analysis for this work, individual MRIs were available for fifteen of the original twenty-three subjects.

Figure 11: Facial Emotion Experimental Stimuli

Experimental stimuli originating from Ekman and Friesen (1976). Hair and non-facial features are hidden, and luminance is controlled. The swirled face and clock are examples of the non-facial control conditions.

MEG Data Preprocessing

The raw, CTF formatted data was converted to the MNE format, and all data preprocessing was accomplished using the MNE software (Gramfort, et al., 2013; Gramfort, et al., 2014). The 275 raw, continuous MEG recordings were bandpass filtered at 1-55 Hz, and signal space projections to remove cardiac and eyeblink contamination were created using automated methods in MNE-Python. Each trial was then isolated from the continuous recordings beginning 50ms pre-stimulus and lasting until 550ms post-stimulus. Data cleaning of these trials was undertaken by thresholding the eyeblink channels as well as the combined MEG recordings and rejecting those trials for which the threshold was exceeded at any point. It was assumed that 20% of all trials collected would be contaminated by either motion artifacts, physiologically generated noise within the subject, or by external electromagnetic interference, and thresholds were set accordingly. This resulted in approximately 700 total trials being retained for each subject from the seven facial conditions (in MNE terminology, these sensor-level individual trials are referred to as *epochs*). Average responses were generated from the cleaned, facial trials of each subject by taking the pointwise mean at each sensor and each time sample across the ensemble of trials (MNE calls this sensor-level average the *evoked response*). The CTF 3rd gradient noise compensation was maintained throughout preprocessing.

A head model based upon the individual MRIs for each subject was created using '*ico-3'* source spacing. Co-registration of the MEG datasets with the head models was accomplished using the MNE graphical user interface (i.e. localization of the subjects' heads within the MEG sensor array). Forward and inverse operators for each subject were created based upon this relationship. *Fixed* source orientations were specified for the forward operators to avoid the issue of discontinuous source timecourses at zero-crossing points. The MEG data of all subjects, including those with individual head models, was also coregistered to the MNE *fsaverage* head model for use in the group model.

Finally, a group average MNE distributed localization (i.e. L2-norm minimum error estimate, also commonly called an MNE "movie") was created using all twenty-three subjects as co-registered to the average head model. Group movies are possible because the morphing operation used here takes place later in the pipeline than the one required by Group-irMxNE, and this group average is used as weighting input by Group-irMxNE. Frames from this group movie are found in Appendix D.

Methods for Facial Emotion Single Subjects Analysis

Although individual analysis of the facial emotion subjects is not a goal of this work, the single subjects were submitted to the analysis pipeline in order to validate the general correctness of their source localizations and spectra prior to inclusion in the group model. The twenty-three subjects, as coregistered to the average brain, were submitted individually to Group-irMxNE (individual subjects analysis may be accomplished by irMxNE but, as an exercise for the newly modified code, they were equivalently submitted one-by-one to the group solver). Inputs required for this submission include the evoked response, noise covariance matrix, forward operator, and the individual-level L2-norm movie. (The movie is used by the solver for two purposes, first to weight the individual source locations and data, and second to exclude from analysis those potential sources which never exceed a specified threshold. The latter function markedly reduces the running time of the solution without any noticeable impact on the outcome.) Group-irMxNE was run repeatedly for each subject in order to tune the regularization parameter to produce eight sources. This allows the individual goodness-of-fit to be compared to the group goodness-of-fit on the basis of the same number of sources. The initial source modeling of each was based upon the window of -50 to 550ms surrounding the facial stimuli.

The average source timecourses output by Group-irMxNE for each subject were then regressed onto the corresponding evoked response resulting in a group of weights for each subject. This regression was based upon the window -50 to 550ms, but the weights were then applied to the sensor-level trials over a window -500 to 500ms. In this way it is possible to construct individual source-level trials for a window which includes the pre-stimulus interval as well as the response using source localizations acquired using only using post-stimulus data. (And therefore, any activity occurring in the pre-stimulus period does not contribute to the localizations.) The thusly-generated source-level trials for each subject were then submitted to the Granger estimation program.

Although no interpretation is performed at the individual level, the localizations were inspected for general fit to the Adolphs model, and their spectra for the occurrence of excessive noise. Although some spectra were noisier than others, ultimately all were included in the group model below.

The fifteen subjects with individual head models were analyzed according to the same procedure. This was done in order to facilitate the comparison between individual and average head models as seen in Table 5. These fifteen subjects are thought to represent the best possible results generated in this work, and their source localizations are seen in Appendix A, and the corresponding conditional Granger spectra in Appendix B.

Methods for Facial Emotion Group Analysis

The twenty-three facial emotion subjects as co-registered onto the average head model were submitted to Group-irMxNE. The inputs required by Group-irMxNE are identical to those of irMxNE, with the exception that lists instead of individual inputs are required for those parameters which are subject-specific (the evoked response, noise covariance matrix, and forward operator). The group L2-norm movie was input to the solver for use as weights. It was decided that eight sources represented the upper limit beyond which both computing resources and difficulty of interpretation became overwhelming. Several iterations of Group-irMxNE were required in order to tune the regularization parameter such that eight sources were present in the output, finally resulting in the group localizations seen in Figure 36.

Group-irMxNE produces one set of source timecourses for each subject (differentiated only by the subject-specific whitening and dewhitening operations). The multivariate multiple regression procedure was used to regress each subject’s source timecourses onto that subject’s evoked response (i.e. the source-level average timecourses were regressed timewise onto the sensor-level average timecourses), resulting in regression weights for each subject. These trials were then input to the Granger estimation program, and the corresponding per-subject conditional Granger influence spectra are seen in Appendix C.

Finally, the per-subject spectra were averaged, resulting in the group conditional Granger influence spectra seen in Figure 37. Goodness-of-fit of the group model to the individual subjects was calculated, and is seen in Tables 5 and 6.

Chapter 3: Results

This chapter contains the results of the work undertaken here. For detailed analysis of these results, see Chapter 4 (Discussion).

Results for Software

The software extensions and new functionality needed to implement the pipeline described above were undertaken, and issues arose which are addressed below.

Group Localization

Group-irMxNE.

The group modifications were initially applied to the original MxNE source code found in MNE-Python (Gramfort, Kowalski, & Hamalainen, 2012). Testing on single subjects and multiple instances of the same subject produced source locations and timecourses identical to those of the unmodified MxNE program. However, when testing reached the single-subject, full-pipeline level of Simulation 3, it was found that connectivity results were either entirely missing or faint, and did not match the expected 40 Hz peak. Since it was known at this point that the multiple regression and Granger causality estimation program worked correctly, suspicion fell upon Group-MxNE. Further testing revealed that the unmodified MxNE program had the same shortcomings, eliminating the possibility of a theoretical or implementation flaw in Group-MxNE. Further testing was performed by attempting to recover the original coefficients of the MVAR system using standard estimation techniques (and a version of the system which did not reverse direction of influence).

. Communications with Alexandre Gramfort and Dennis Strohmeier of the MNE development team (4/20/2015) revealed that a newer solver in the MxNE family, irMxNE had recently been released in order to address issues of poor source localization and biased amplitude, and it was suggested that the group modifications be ported to the new code (Strohmeier, 2014; Strohmeier, Haueisen, & Gramfort, 2014). This was done and Group-irMxNE immediately began to produce correct results in the full pipeline test of Simulation 3.

irMxNE single trials regression.

The multivariate multiple regression procedure was implemented in a Python function and seen to function well. A requirement of multivariate multiple regression is that there exist more observations than predictors. During testing it was accidentally discovered that the implementation in the Statsmodel package used here is able to function without this requirement being met (Seabold & Perktold, 2010). It is anticipated that the function is able to work around the violation of this requirement due to high levels of multicollinearity among the predictors. The normally output goodness of fit measurements are not available from the regression function, however.

Connectivity

Granger causality.

Wavelet domain Granger causality was initially implemented in a custom Matlab function based on details provided in the publications of Dhamala, Rangarajan, and Ding (2008a; 2008b). This function continued to use Dhamala and Rangarajan's implementation of the Wilson spectral matrix decomposition found in the Fieldtrip package (Matlab function *sfactorization_wilson*). This code performed well and the exercise was informative. Conditional Granger causality is considerably more complex however, and would have been undesirable to implement by hand. The Fieldtrip implementation of Granger and conditional Granger causality is undocumented, but was found after considerable searching. Furthermore, invoking this code required understanding the complex control structures necessary to access it. This was accomplished and the time-frequency domain conditional Granger functionality was implemented as described in the previous methods.

Testing of the connectivity functionality also proceeded as described. The MVAR system of Dhamala, Rangarajan, and Ding is submitted directly to this function in the first part of Simulation 3, and is utilized in testing the full pipeline in Simulations 3-6. Initial testing revealed a slight skew of the peak of the Granger spectra towards lower frequencies. The peak frequency described by Dhamala, Rangarajan, and Ding is approximately 40 Hz, but a peak closer to 35 Hz was seen. In personal communications, Mukesh Dhamala (October, 2013) suggested using a Morlet wavelet with parameter w_0=12 instead of the more common w_0=6. This change biases the results towards precise estimation of frequency at the expense of precision in the time domain (Torrence &

Compo, 1998). Although this change minimized the skew being seen, it continues to be present, but is not judged to be detrimental to interpretation of the results in this work. A discussion and demonstration of this skew is presented in Appendix E.

Goodness of Fit

The goodness-of-fit procedure suggested by Alexandre Gramfort (personal communications, October, 2014) based on correlation of serialized sensor-space data was implemented and appears to work well. Percentages of variance accounted for range from almost 100% for simulations with very low levels of noise added, to just above 0% in one of the facial emotion subjects in the group model (Table 5).

Inference

Single-subject, resampling-based inference following the model of Barnett and Seth (2014b; 2010) was implemented as part of the Granger causality program in Matlab. Results are seen to be good, although this type of inference makes intensive use of both computer processor and memory resources. For example, for spectra containing 60 frequencies, 300 time samples, and four sources, and running four parallel threads, it takes approximately eight hours to complete 250 reorderings on a workstation class computer. A maximum of 2,500 reorderings is possible with the current code (which keeps all results in memory at once) before 32 GB of memory are exceeded. In the absence of an analytic method for performing inference, it is important that this algorithm be optimized to the fullest extent possible.

Results for Simulations

Results for Simulation 1

The pair of time-domain signals created for this simulation are seen in Figure 12, their Fourier power and cross power in Figure 13, and their time-frequency wavelet power and cross power in Figure 15. Figure 14 shows the data in Figure 15 averaged over the time axis to enable comparison with the Fourier spectra.

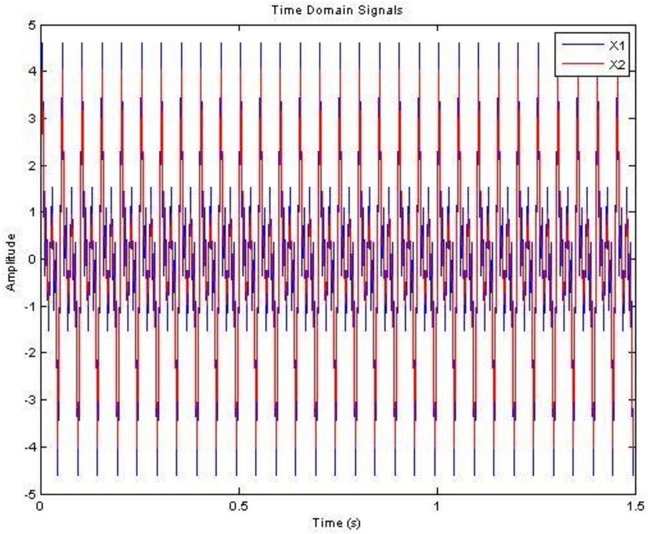


Figure 12: Simulation 1 Time Signals

Time domain signals x_1 and x_2 created by summing several pure sinusoids. Although the contributions of the individual sinusoids are difficult to discern, periodic components are clearly visible in the signals.

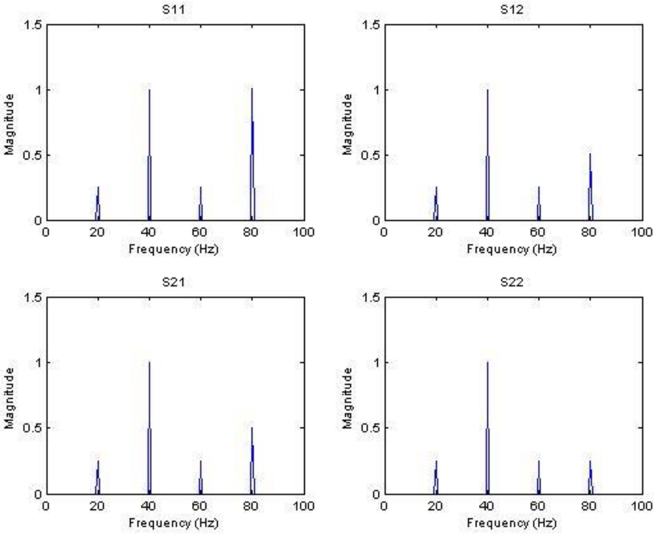


Figure 13: Simulation 1 Fourier Power and Cross Power

Magnitude of the Fourier power and cross power spectra (magnitude is used because the cross spectra are complex-valued). The magnitude power spectra for x_1 and x_2 are seen in the upper left and lower right panels respectively. The magnitudes cross power spectra are seen in the lower left and upper right. The cross power are complex conjugates of one another, but the magnitude operation obscures this. In x_1 the 40 and 80 Hz waves are twice the amplitude of those at 20 and 60 Hz. Since power represents a squaring operation, these waves are seen to have four times the magnitude in the upper left pane. A similar effect is seen for x_2 in the lower right. The cross power are simply the pointwise products of x_1 and conjugate(x_2) and vice versa.

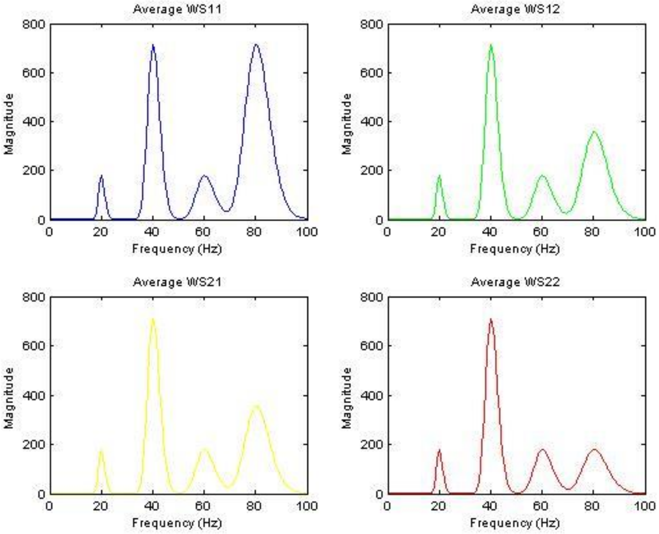


Figure 14: Simulation 1 Time Average of Wavelet Power and Cross Power

Average wavelet power and cross power spectra in time. The panels from the next figure are seen here averaged over the time axis in order to facilitate comparison with the Fourier spectra above. The magnitude peaks remain constant per the scaling performed. The scaled wavelets become wider as frequency increases and the wavelet spectra does not drop to zero between peaks at higher frequencies. This effect is particularly evident between the 60 and 80 Hz peaks in the lower right pane.

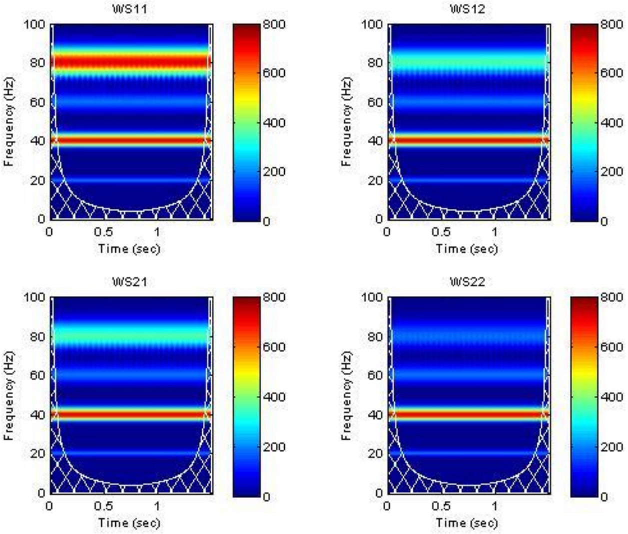


Figure 15: Simulation 1 Wavelet Power and Cross Power

Magnitude of the wavelet power and cross power spectra for the signals created in Simulation 1.Time is seen on the x-axes and frequency on the y-axes. Magnitude is shown according to the colorbars at the right of each pane. Since the system does not change in time the frequencies remain constant from left to right.

Results for Simulation 2

The two individual waveforms drawn from median nerve stimulation timecourses are seen in Figure 16. The 500 trials created by varying the length and amplitude of these waveforms are seen in Figures 17 and 18. Wavelet power and cross power are seen in Figure 19, a), the conditional Granger spectra in Figure 19, c), and resampling-based inference in Figure 19, b) and d).

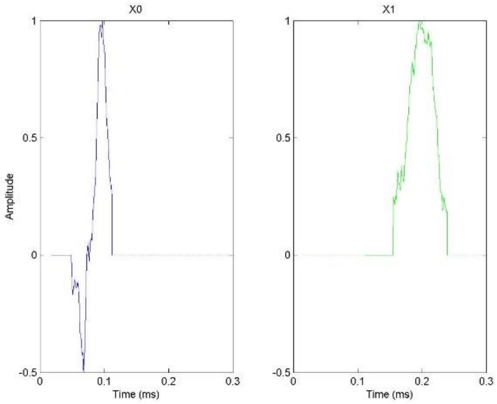


Figure 16: Simulation 2 Components

Two source waveforms, x_0 and x_1, taken from a median nerve stimulation analysis. Note that the time axes here are condensed so the waves may be viewed more easily, and do not match those of subsequent Figures 17, 18, and 19.

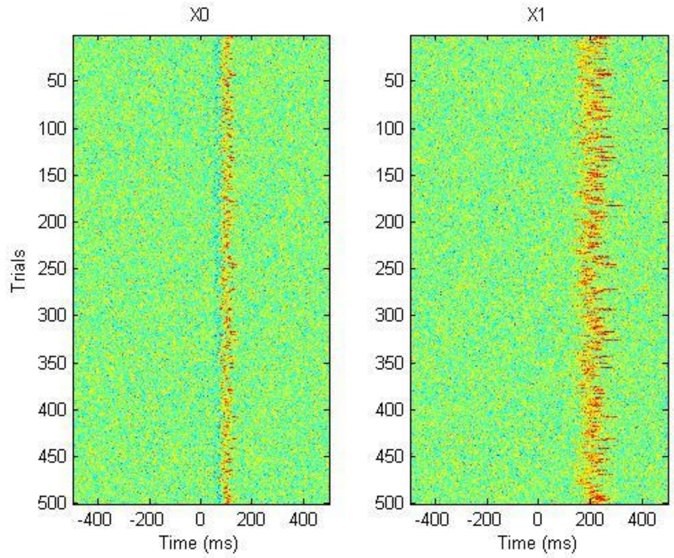


Figure 17: Simulation 2 Trials Flat Plot

Flat plots of individual trials. Time is shown on the horizontal axis, trials on the vertical, and signal amplitude is represented with color. The waveforms themselves are scaled in both amplitude and duration, and all four scaling factors (x_0 amplitude, x_0 duration, x_1 amplitude, x_1 duration) are correlated at r=0.8. The second waveform begins immediately following completion of the first. In other words, the amplitude and duration of the first component predict the starting time, amplitude, and duration of the second. This is intended to approximate different strengths and latencies of neural response, with x_0 causing x_1. Sampling is taken to be 200 Hz.

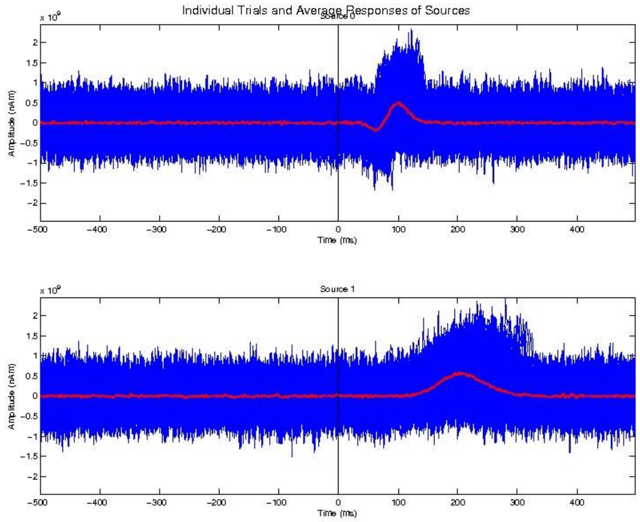


Figure 18: Simulation 2 Trials and Trials Average

Another view of the individual trials shown in Figure 17. Individual trials are over-plotted in blue, and their time-wise average is shown in red Waveform x_0 begins at time 0, followed immediately by x1. Note that a pre-stimulus period is included to which the post-stimulus response may be compared (an informal null hypothesis).

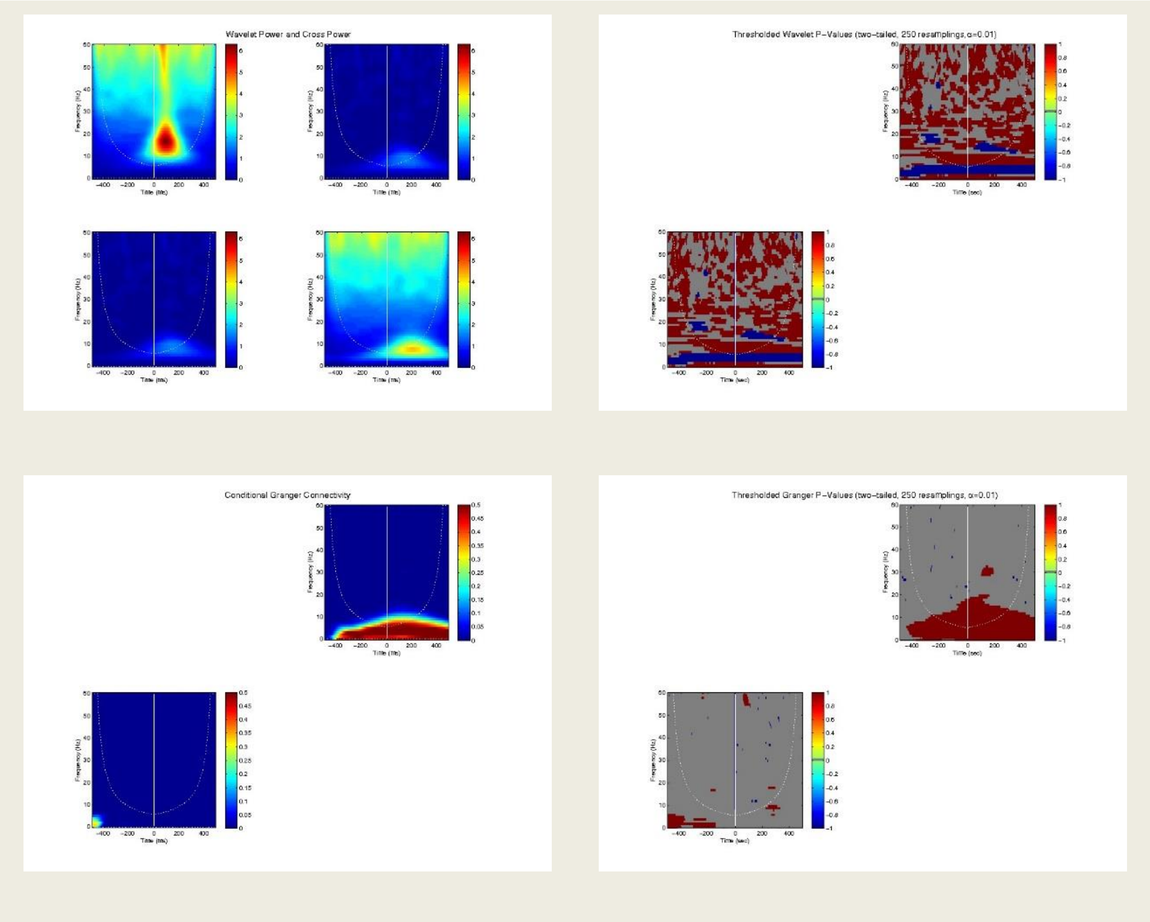


Figure 19: Simulation 2 Source Power and Granger Spectra

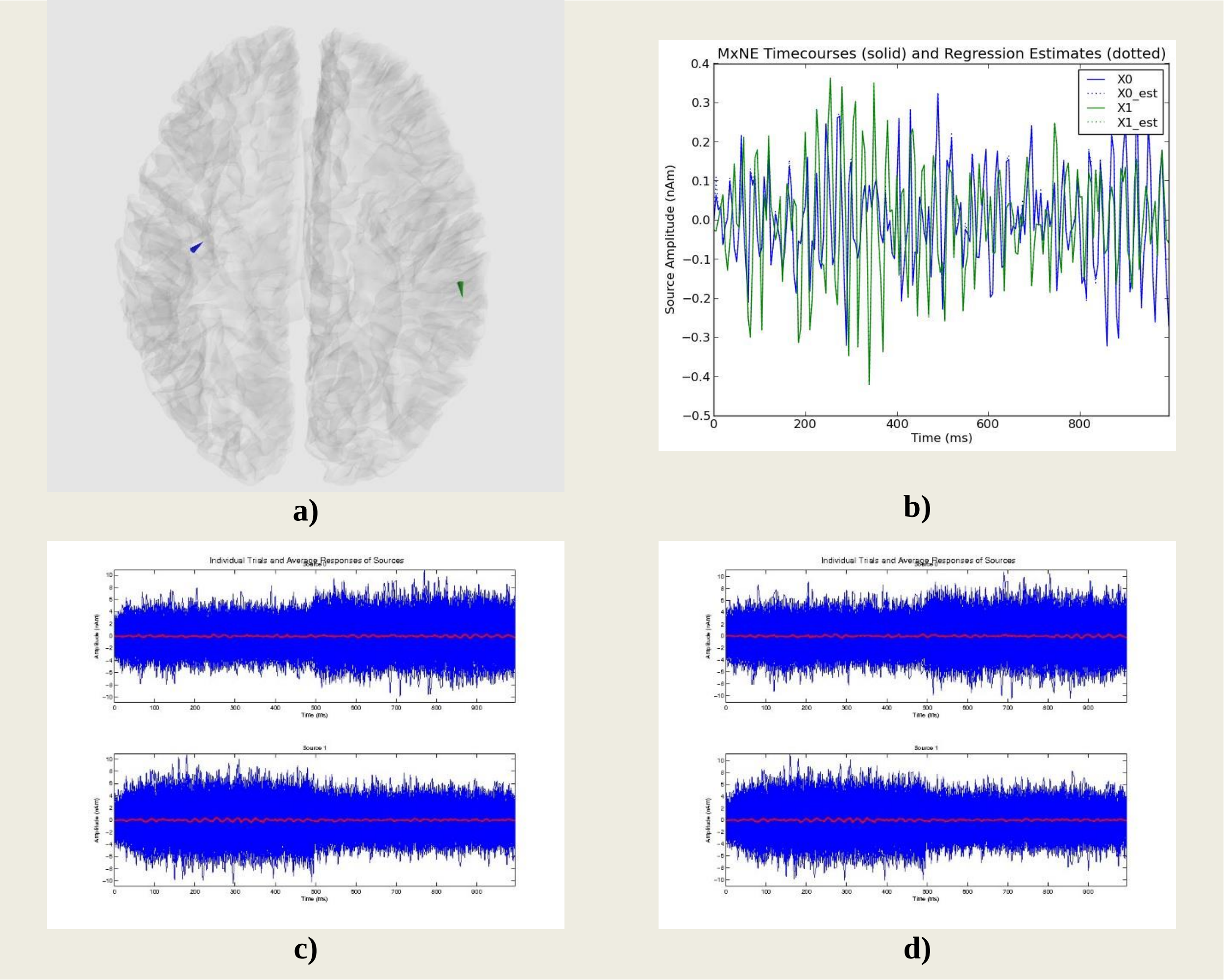
Power a) and conditional Granger spectra c), along with inference for each, b) and d). These spectra are derived from the 500 trials seen in Figures 17 and 18, of the components in Figure 16. The onset of power and cross power follows the stimulus onset (vertical line), but are also seen to bleed backwards into the pre-stimulus period. The Granger spectra clearly shows source x_0 to be causing x_1 (panel c), upper right) but not so in the opposite direction (panel c) lower left). 250 reorderings of resampling-based inference utilizing a two-tailed test with α=0.01 show the conditional Granger causality to be significant, but the wavelet cross-power is inconclusive. Panels b) and d) are constructed such that locations where the spectra a) and c) are significantly greater than the null spectra are shown in red, and less than the null spectra in blue.

Results for Simulation 3

The original source localizations in the left and right auditory cortices are shown in Figure 20, a). Group-irMxNE returned this same pair of locations. The 300 simulated input trials are overplotted in blue and their average in red in panel c), with x_0 seen in the top and x_1 in the bottom. The average source timecourses returned by Group-irMxNE are labeled ‘X0’ and ‘X1’ in solid lines in b). The individual source trials generated by the multivariate multiple regression are overplotted in blue in d), and their averages are seen in red in d) as well as in dotted lines labeled ‘X0_est’ and ‘X1_est’ in b). In b) ‘X0’ and ‘X0_est’ and ’X1’ and ‘X1_est’ may be compared to judge the success of the regression (the dotted lines are obscured behind the solid lines in this figure). See Figure 9 for an overview of the pipeline, and Figure 10 for a diagram of the data processing in this simulation.

The wavelet power and cross power, Granger, and conditional Granger spectra of the original trials are seen in the left-hand column of Figure 21, in panels a), c), and e). The wavelet power spectra of x_0 and x_1 are seen in the top left and bottom right corners of panel a) respectively. The cross spectra of x_0 and x_1 are seen in the bottom left and top right of the same panel. These spectra correspond to the originally simulated trials shown in panel c) of Figure 20. The Granger connectivity spectra of x_0 and x_1 is seen in panel c) of Figure 20. Since Granger connectivity is only defined between series, only the bottom left and top right portions of this panel are seen. The top right contains the Granger spectra $x_0 \rightarrow x_1$, and the bottom left $x_1 \rightarrow x_0$ (row-causing-column). Panel e) contains the conditional Granger connectivity. The right-hand column of Figure 21 shows the resampling-based inference performed on the wavelet power b), Granger d), and conditional Granger spectra f) respectively. 250 reordering are undertaken, and a two-tailed test with $\alpha=0.01$ is used. Those spectral locations which are significantly greater than the null are shown in red, and those significantly less in blue. The panels in Figure 22 correspond directly to those in Figure 21, but are based on the trials output from the Group-irMxNE rather than those input to it.

Table 1 shows the goodness of fit results for Simulation 3.



a)

b)

c)

d)

Figure 20: Simulation 3 Source Localizations and Timecourses

Group-irMxNE source localizations a), source timecourses b) (solid lines), averaged individual trial timecourses b) (dotted lines, obscured by solid) and d) (red), and over-plotted individual trial timecourses d) (blue). Localizations a) may be matched to their corresponding timecourses by color b). Indexing in the legend of b) matches indexing from the top in c) and d). The original simulated trials are seen in c) and may be compared to the trials recovered by the pipeline d). The simulated trials are constructed from Equations M and N and, in the first half of the time window, positive influence $x_0 \rightarrow x_1$ is seen as increased variance in x_1 versus x_0. This pattern reverses in the second half.

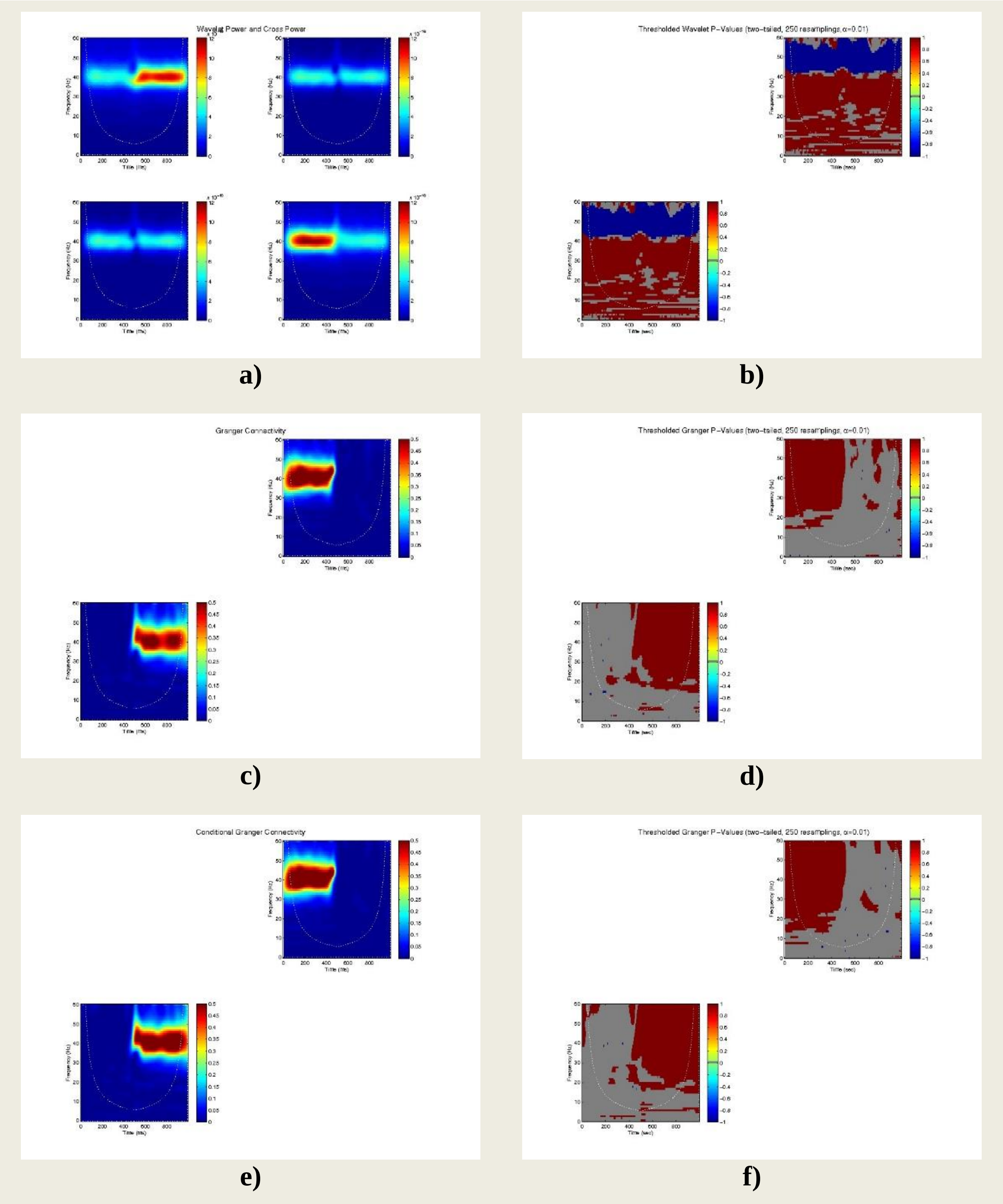
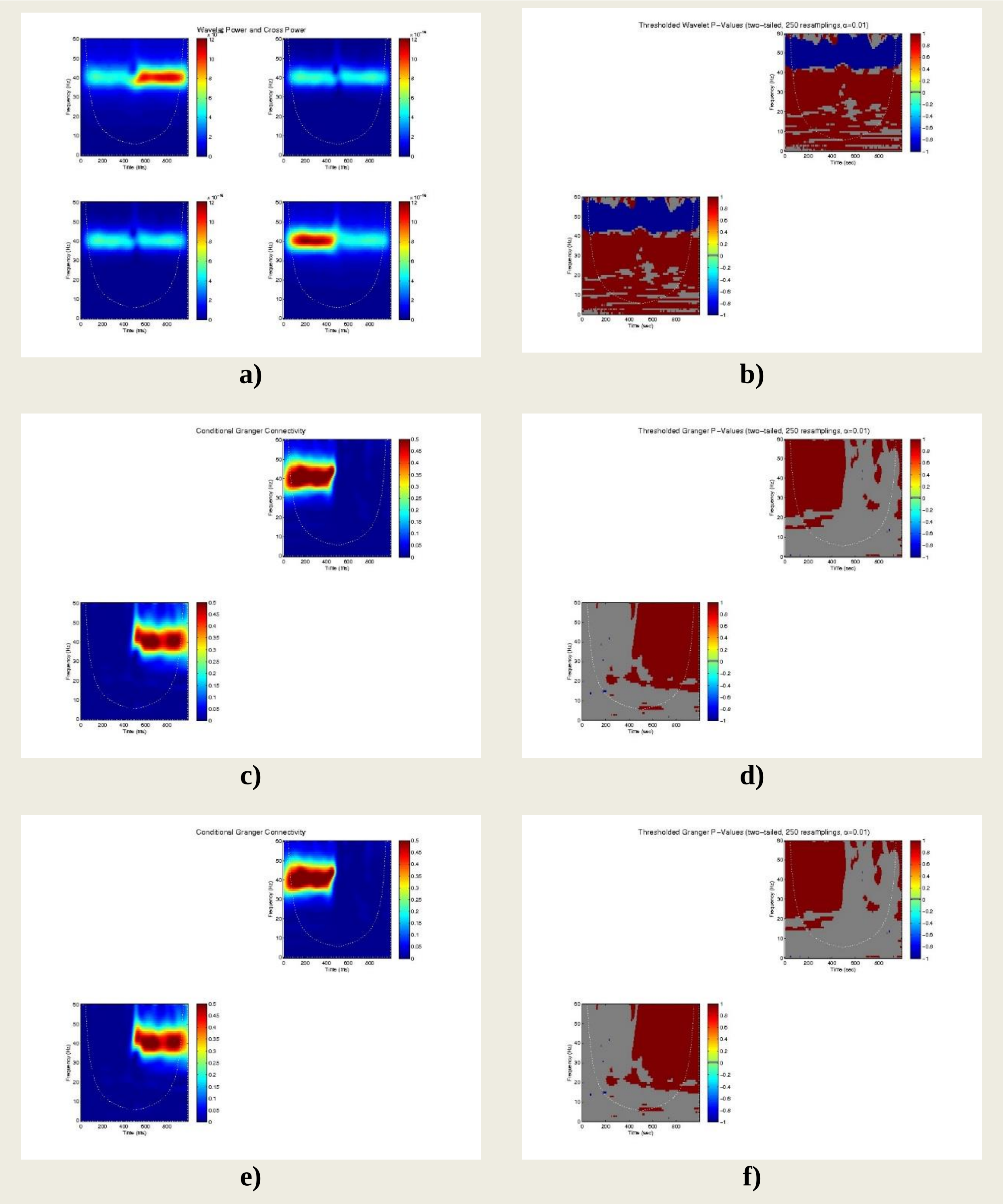


Figure 21: Simulation 3 Original Source Power and Granger Spectra

Source timecourse power and cross power of the original simulated trials a), Granger connectivity c), conditional Granger connectivity e), and inference b), d), and f). This data corresponds to the trials in Figure 20, c). The pattern of variance in the power spectra a) matches that of the trials in Figure 20, c) and d). Granger and conditional Granger spectra indicate causality $x_0 \rightarrow x_1$ during the first half of the time window, reversing to $x_1 \rightarrow x_0$ in the second (remember the convention of "row-causing-column"). This pattern is supported by inference results, although inference on the original power and cross power spectra is again inconclusive.



a) **b)**

c) **d)**

e) **f)**

Figure 22: Simulation 3 Group-irMxNE Source Power and Granger Spectra

Source timecourse power and cross-power of the trials at the output of Group-irMxNE and the multivariate regression a), Granger connectivity d), conditional Granger connectivity e), and inference b), d), and f). This figure may be compared directly to the results based on the original simulated data in Figure 21. These results indicate that the pipeline constructed here is able to recover not only the locations of the simulated sources, but also their timecourses with sufficient fidelity to reproduce the original pattern of connectivity.

Subject	irMxNE Total ('r' #1)	Src 0 ('r' #1)	Src 1 ('r' #1)	Trials Total ('r' #2)	Src 0 ('r' #2)	Src 1 ('r' #2)
0	**0.84 (0.71)**	**0.27 (0.07)**	**0.80 (0.65)**	**0.84 (0.71)**	**0.27 (0.07)**	**0.80 (0.65)**

Table 1: Simulation 3 Goodness of Fit

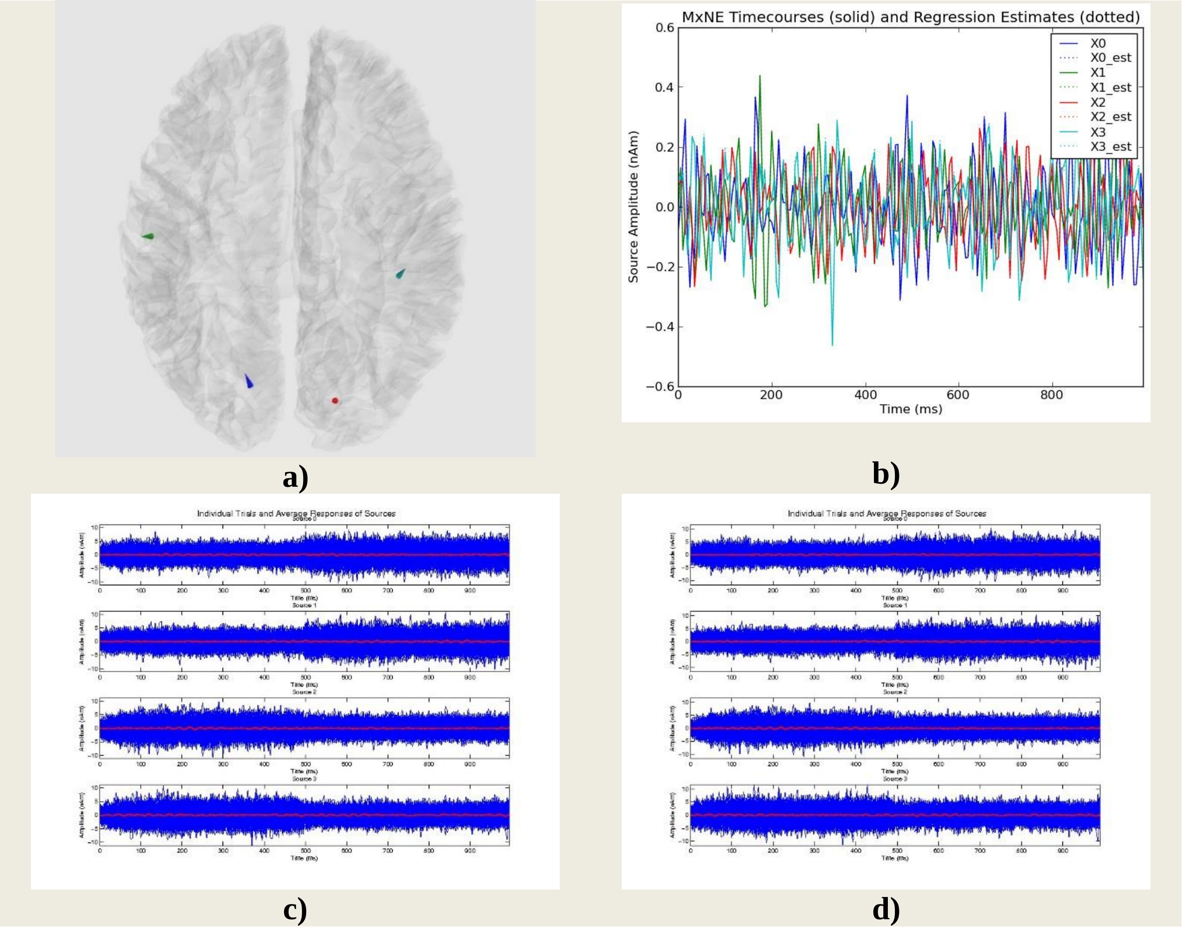
Goodness of fit results for the Group-irMxNE model (columns 2-4) and Regression Trials Model (columns 5-7). See Figure 10 (red, bottom right) for the location of these results in the pipeline. Values presented here are correlations and squared correlations, which are interpreted as percentages of variance. The format is, r (r^2). Inference is performed at α=1e-6, and significance is indicated in boldface. The average timecourses output from Group-irMxNE and translated back to sensor-space via the forward operator are seen to account for 71% of the variance of the original evoked input (column 2). The two individual sources account for 7% and 65% of the variance respectively (columns 3, 4). Instead of using the timecourses output from Group-irMxNE, the Regression Trials Model translates the individual trials output from the multivariate regression back to sensor-space and averages them (Figure 9). This model subsumes the Group-irMxNE functionality, and adds the multivariate regression. Results in columns 5-7 match those of 2-4, indicating that the multivariate multiple regression does not degrade the data to any measurable extent.

Results for Simulation 4

The original pairs of source localizations in the left and right auditory cortices and left and right primary visual areas are shown in Figure 23, a). Group-irMxNE returned these same pairs of locations. The simulated input trials are overplotted in blue and their average in red in pannel c) (x_0: left visual, blue, x_1, left auditory, green, x_2: right visual, red, x_3: right auditory, cyan). The average source timecourses returned by Group-irMxNE are labeled ‘X0’ through ‘X3’ in solid lines in b). The individual source trials generated by the multivariate multiple regression are overplotted in blue in d), and their averages are seen in red in d) as well as in dotted lines labeled ‘X0_est’ through ‘X3_est’ in b). In b) ‘X0’ and ‘X0_est’, etc. may be compared to judge the success of the regression (the dotted lines are again obscured behind the solid lines in this figure).

The wavelet power and cross power, Granger, and conditional Granger spectra of the original trials are seen in the left-hand column of Figure 24, in panels a), c), and e). The wavelet power spectra of x_0 through x_3 are seen along the diagonal of panel a). Non-diagonal elements represent the cross spectra, index by row and column from the top left. These spectra correspond to the originally simulated trials shown in panel c) of Figure 23. The Granger connectivity spectra of x_0 through x_3 is seen in panel c) of Figure 24. Panel e) contains the conditional Granger connectivity. The right-hand column of Figure 24 shows the resampling-based inference performed on the wavelet power b), Granger d), and conditional Granger spectra f) respectively. The panels in Figure 24 correspond directly to those in Figure 25, but are based on the trials output from the Group-irMxNE rather than those input to it.

Table 2 shows the goodness of fit results for Simulation 4.



a)

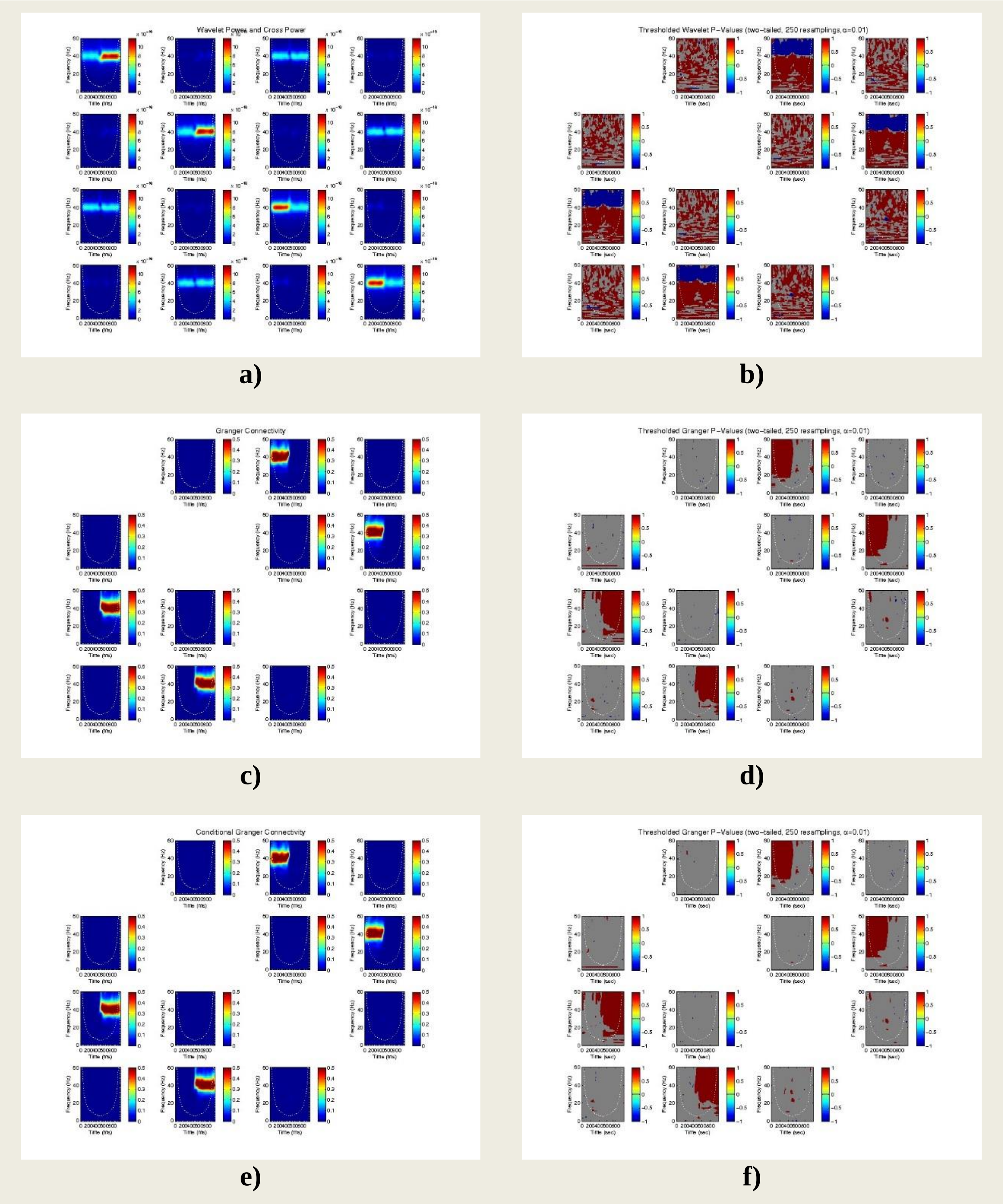
b)

c)

d)

Figure 23: Simulation 4 Source Localizations and Timecourses

Group-irMxNE localizations a), timecourses b), trials d), and originally simulated data c). These results may be interpreted similarly to those of Simulation 3 in Figure 20, except that the subject here contains two pairs of linked sources (left and right auditory, left and right visual). The source activity consists of two instances of the system used in Simulation 3, with the left and right auditory being represented by sources ‘0’ and ‘2’, and the visual by ‘1’ and ‘3’. The two systems are independent of one another.



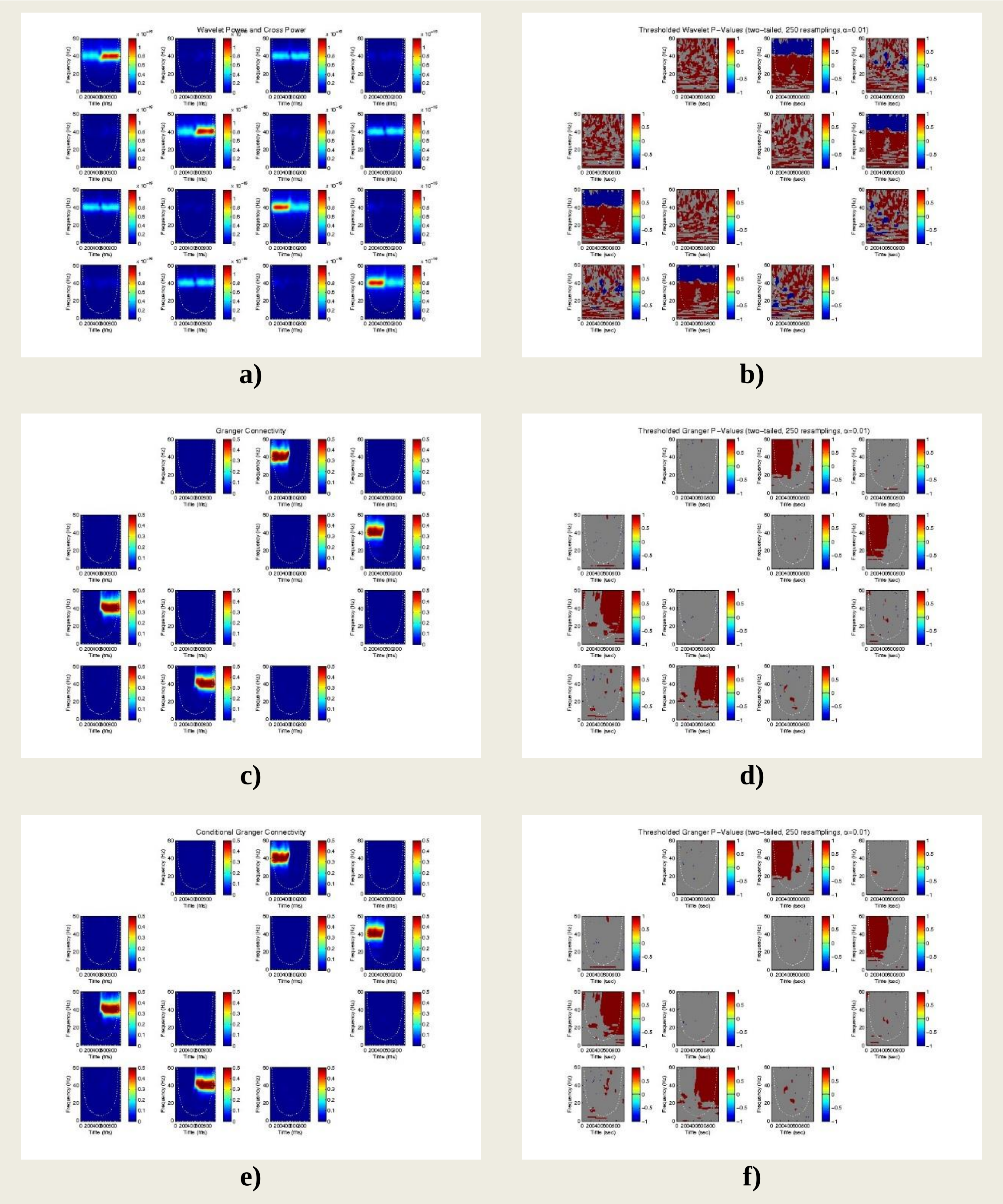
a) **b)**

c) **d)**

e) **f)**

Figure 24: Simulation 4 Original Power and Granger Spectra

Source timecourse power and cross power of the original simulated trials a), Granger connectivity c), conditional Granger connectivity e), and inference b), d), and f). This data corresponds to the trials in Figure 24, c), and this figure is interpreted in a manner similar to Figure 21 in the previous simulation. The pattern of variance in the power spectra a) matches that of the trials in Figure 24, c) and d). Granger and conditional Granger spectra indicate causality $x_0 \rightarrow x_2$ and $x_1 \rightarrow x_3$ during the first half of the time window, reversing to $x_2 \rightarrow x_0$ and $x_3 \rightarrow x_1$ in the second. This pattern is supported by inference results.



a) **b)**

c) **d)**

e) **f)**

Figure 25: Simulation 4 Group-irMxNE Power and Granger Spectra

Source timecourse power and cross-power of the trials at the output of Group-irMxNE and the multivariate regression a), Granger connectivity d), conditional Granger connectivity e), and inference b), d), and f). This figure may be compared directly to the results based on the original simulated data in Figure 24. These results indicate that the pipeline constructed here is able to recover the pattern of connectivity between two pairs of simulated sources.

Subject	Trials Total	Src 0	Src 1	Src 2	Src 3
0	**0.82** **(0.67)**	**0.36** **(0.13)**	**0.44** **(0.20)**	**0.45** **(0.21)**	**0.36** **(0.13)**

Table 2: Simulation 4 Goodness of Fit

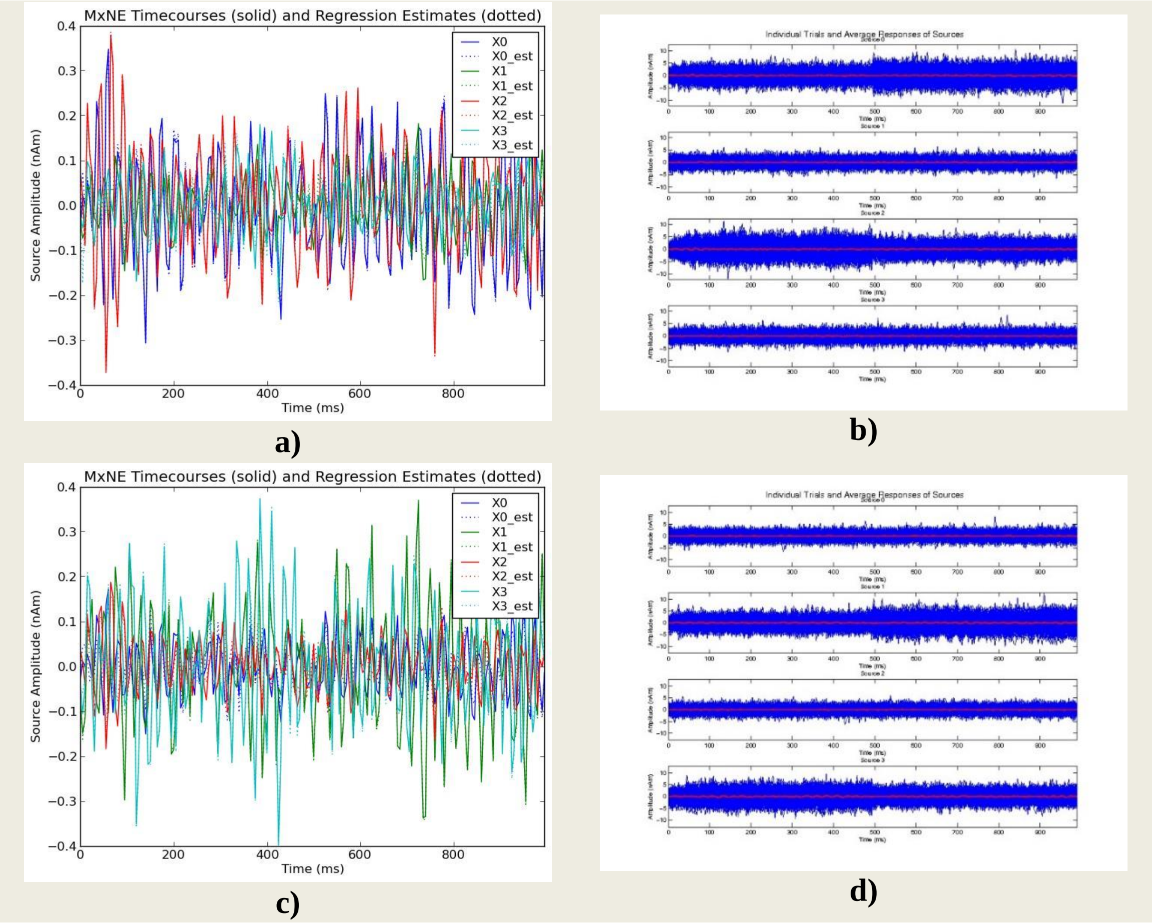
Goodness of fit results for the Regression Trials Model (since the Group-irMxNE model is subsumed by the Regression Trials Model, it is not presented). The combined model is seen to account for 67% of the variance of the original evoked data. Individual sources account for 13%, 20%, 21%, and 13%, respectively.

Results for Simulation 5

The original pairs of source localizations in the left and right auditory cortices of subject ‘0’ and left and right primary visual areas of subject ‘1’ correspond to those shown in Figure 23, a). Group-irMxNE returns these same pairs of locations. The input data matches that seen in panel c) of the same figure. The average source timecourses returned by Group-irMxNE for subject ‘0’ are labeled ‘X0’ through ‘X3’ in solid lines in Figure 26, a), and for subject ‘1’ in c). The individual source trials generated by the multivariate multiple regressions are overplotted in blue in b) and d), and their averages are seen in red. These averages are also seen in dotted lines in a) and c) as well.

The wavelet power and cross power and conditional Granger spectra of the regressed trials for subject ‘0’ are seen in the left-hand column of Figure 27, in panels a) and c), and for subject ‘1’ in Figure 28 a) and c). Figure 29 shows the group average of the wavelet a) and conditional Granger spectra b) for subjects ‘0’ and ‘1’, although no group inference is provided. These spectra may be directly compared to the spectral results of Simulation 4, Figures 24 and 25, panels a) and e). The color axis scaling is equivalent in the Granger panels.

Table 3 shows the goodness of fit results for Simulation 5.



a)

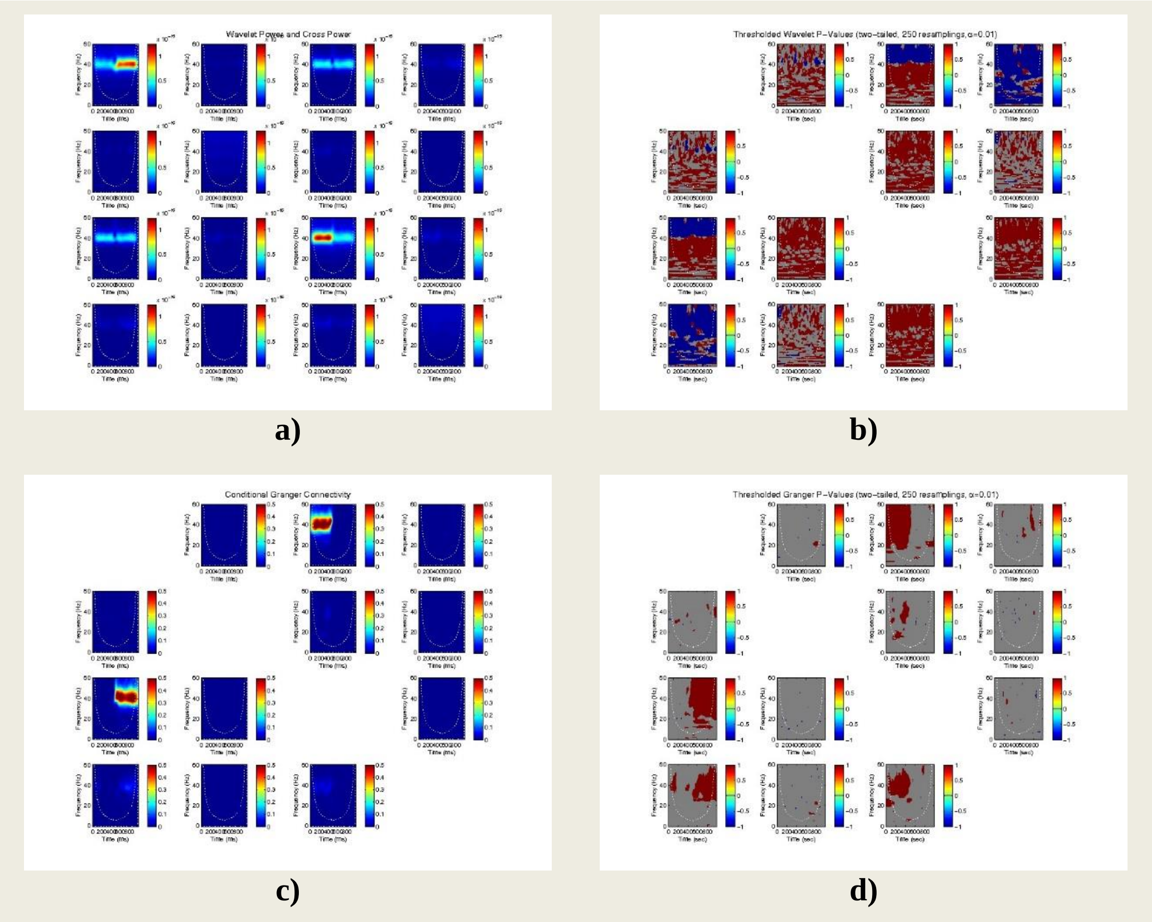
b)

c)

d)

Figure 26: Simulation 5 irMxNE Source Timecourses and Trials

Group-irMxNE timecourses, and individual trials for subjects ‘0‘and ‘1’. This simulation reuses the localizations and pair of interacting sources from Simulation 4 (Figure 23) but, in this case, the pairs of sources are assigned to two separate subjects. The result is four localizations which match Figure 23, a). Since this output represents a group model, each subject contains four sources, but only two are active in each. The Group-irMxNE average timecourses for subject ‘0’ are seen in a), and the resulting individual trials for this subject in b). Corresponding output for subject ‘1’ is seen in c) and d). Notice that b) contains the pattern of interacting sources seen in previous simulations in sources ‘0’ and ‘2’, while sources ‘1’ and ‘3’ display a constant level of noise throughout.

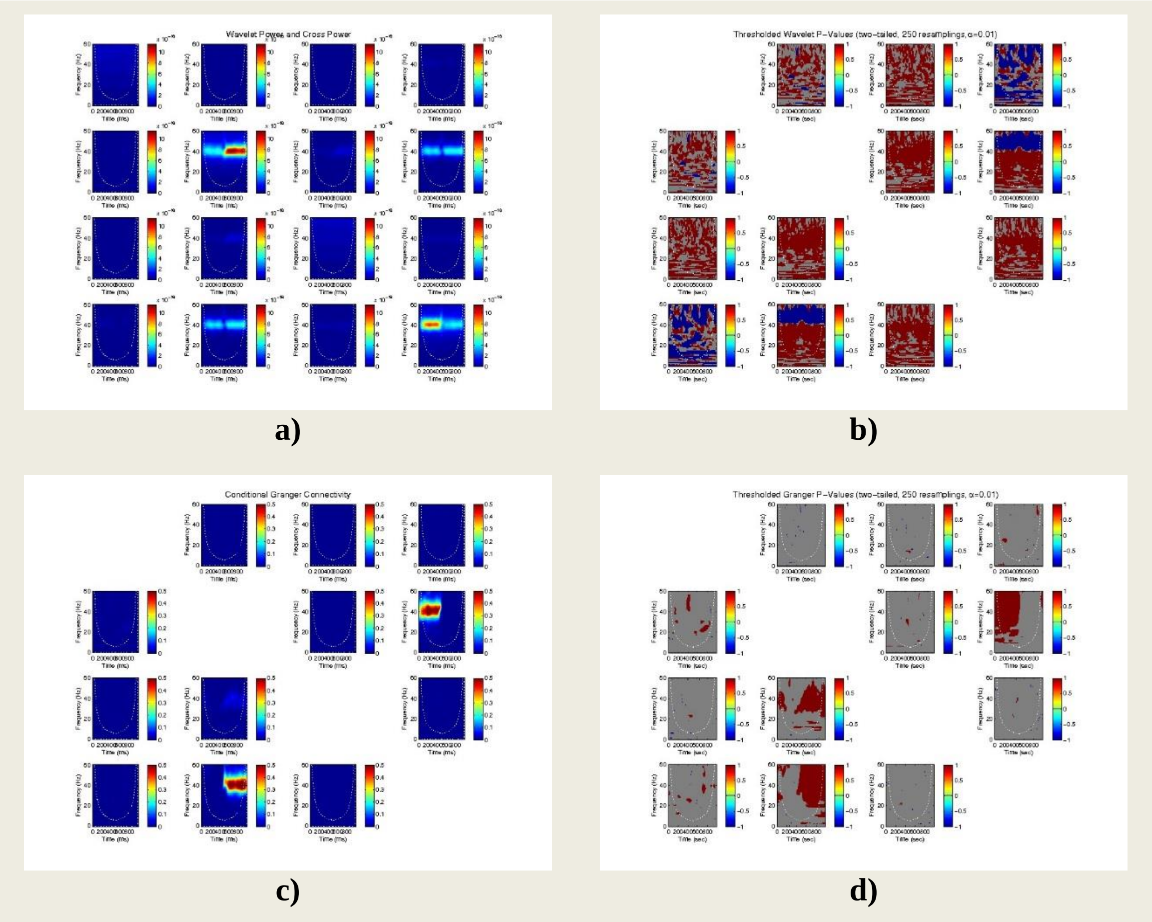


a) **b)**

c) **d)**

Figure 27: Simulation 5 Subject 0 Power and Granger Spectra

Source timecourse power and cross-power for subject ‘0’ a), conditional Granger connectivity c), and inference b), and d). Since Granger and conditional Granger spectra are seen to be in good agreement in previous simulations, only the later are presented here. These results correspond to the individual trials output from Group-irMxNE and the multivariate regression as seen in Figure 26, b). The Granger spectra indicates the reversing pattern of connectivity between sources ‘0’ and ‘2’, representing the pair of sources assigned to this subject. Hints of this activity are also seen with source ‘3’ influencing sources ‘0’ and ‘2’. Generally these results indicate that the pipeline is able to recreate the originally simulated activity for the subject.

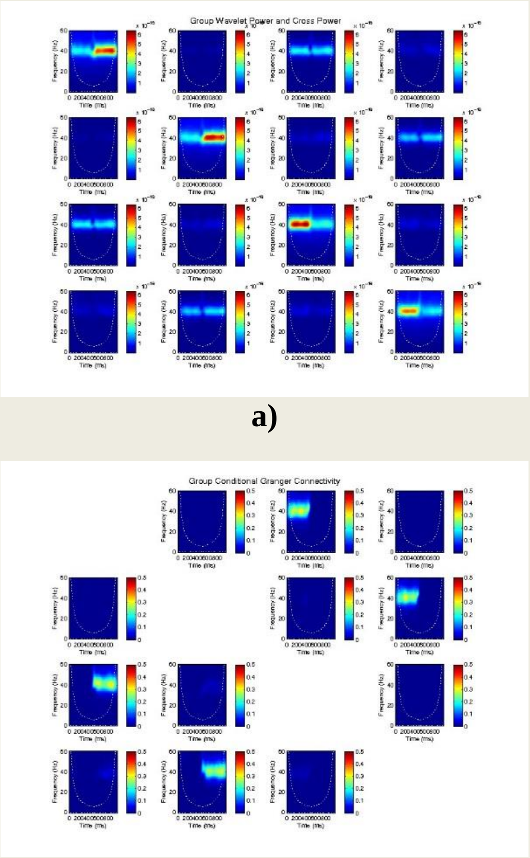


a) **b)**

c) **d)**

Figure 28: Simulation 5 Subject 1 Power and Granger Spectra

Source timecourse power and cross-power for subject ‘1’ a), conditional Granger connectivity c), and inference b), and d). These results correspond to the individual trials output from Group-irMxNE and the multivariate regression as seen in Figure 26, d). The Granger spectra indicates the reversing pattern of connectivity between sources ‘1’ and ‘3’, representing the pair of sources assigned to this subject. Hints of this activity are also seen with source ‘2’ influencing sources 12’. Generally these results indicate that the pipeline is able to recreate the originally simulated activity for the subject.



a)

b)

Figure 29: Simulation 5 Group Power and Granger Spectra

Mean source timecourse power and cross-power a) and conditional Granger connectivity b). These spectra are the simple means of those in Figures 27 and 28, a) and c). In combination with the localizations seen in Figure 23, a), they represent the group model for the pair of subjects entered into this simulation. Inference is not provided since a group procedure has not been identified. These results may be compared to panels a) and c) of Figure 25, where the same pattern of activity is analyzed within a single subject. The color axis of the power spectra is scaled according to the maximum value in both figures, and therefore is not equal. The color axis is equalized for the conditional Granger spectra however, and diminished power is seen here versus Figure 25, c). Overall these results show the group model to faithfully represent a combination of the two subjects, and that the original pairs of sources remain distinct. Since the source activity of the two subjects differs, they do not represent good candidates for a group model, and these results reflect the expected outcome of combining two disparate entities.

Subject	Total	Src 0	Src 1	Src 2	Src 3
0	**0.73** **(0.54)**	**0.50** **(0.25)**	-0.02 (0.00)	**0.64** **(0.41)**	**0.06** **(0.00)**
1	**0.76** **(0.58)**	**0.05** **(0.00)**	**0.69** **(0.48)**	**-0.02** **(0.00)**	**0.45** **(0.20)**

Table 3: Simulation 5 Goodness of Fit

Goodness of fit results for the Regression Trials Model. As opposed to previous simulations, two subjects are present in these results, and goodness of fit is assessed separately for each subject. Fit is based on individual subjects' trials propagated through the group localizations, and these values represent the fit of the group localization to each subject. (It may be useful to reference Figures 9 and 10.) This model is seen to account for 54% of the variance of the original evoked response of subject '0', and 58% for subject '1'. Fit of individual sources indicates which sources best explain the activity of which subjects. It is seen that, for subject '0', sources '0' and '2' represent 25% and 41% variance, while sources '1' and '3' contribute very little. The opposite pattern is seen for subject '1'.

Results for Simulation 6

The original pairs of source localizations in the left and right auditory cortices of three subjects are seen in Figure 30, a) (the localizations for the three subjects are artificially combined in this panel). The localizations returned by Group-irMxNE are seen in b). The average source timecourses returned by Group-irMxNE for the three subjects are seen in Figure 31, a), c), and e). The individual source trials generated by the multivariate multiple regressions are overplotted in blue in b), d), and f), and their averages are seen in red. These averages are also seen in dotted lines in a), c), and e). The regression-estimated trial averages seen in dotted lines are no longer completely obscured by the solid lines.

The wavelet power and cross power and conditional Granger spectra of the regressed trials for the three subjects are seen in Figures 32, 33, and 34. Figure 35 shows the group average of the wavelet a) and conditional Granger spectra b) for the three subjects. The color axis scaling is equivalent in the Granger panels.

Table 4 shows the goodness of fit results for Simulation 6.

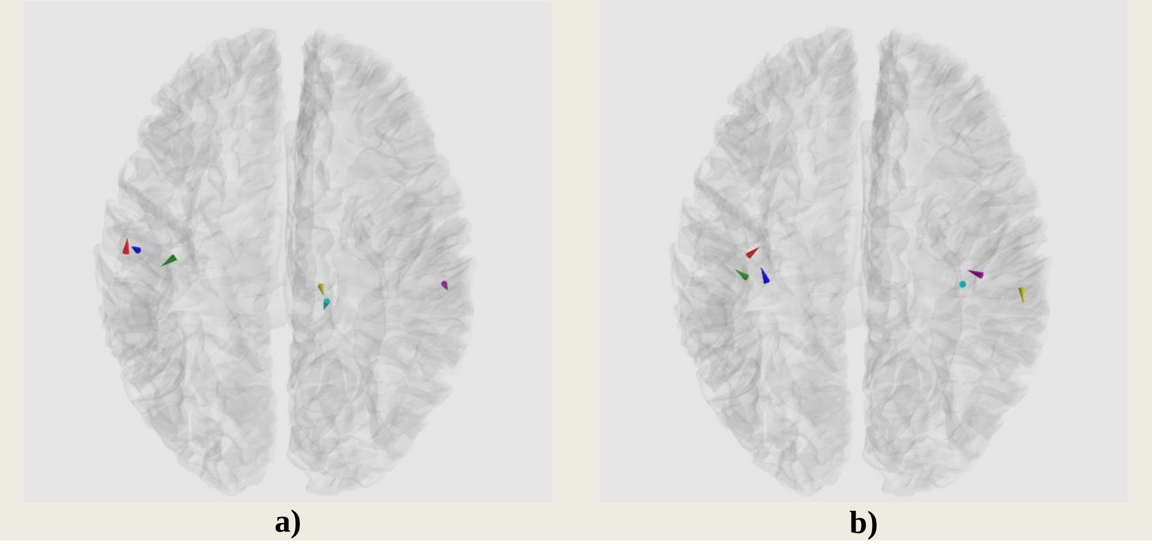
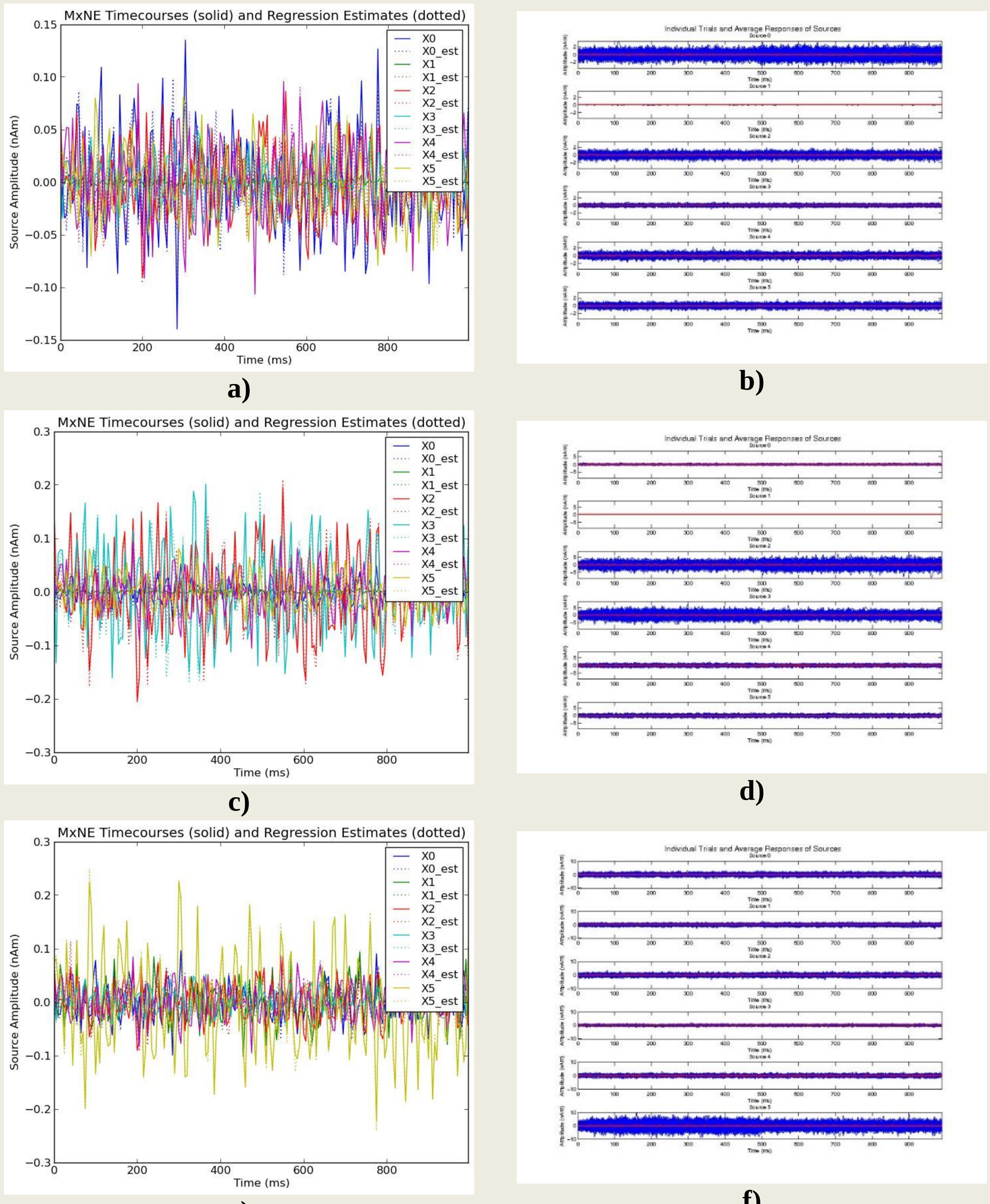


Figure 30: Simulation 6 Source Localizations

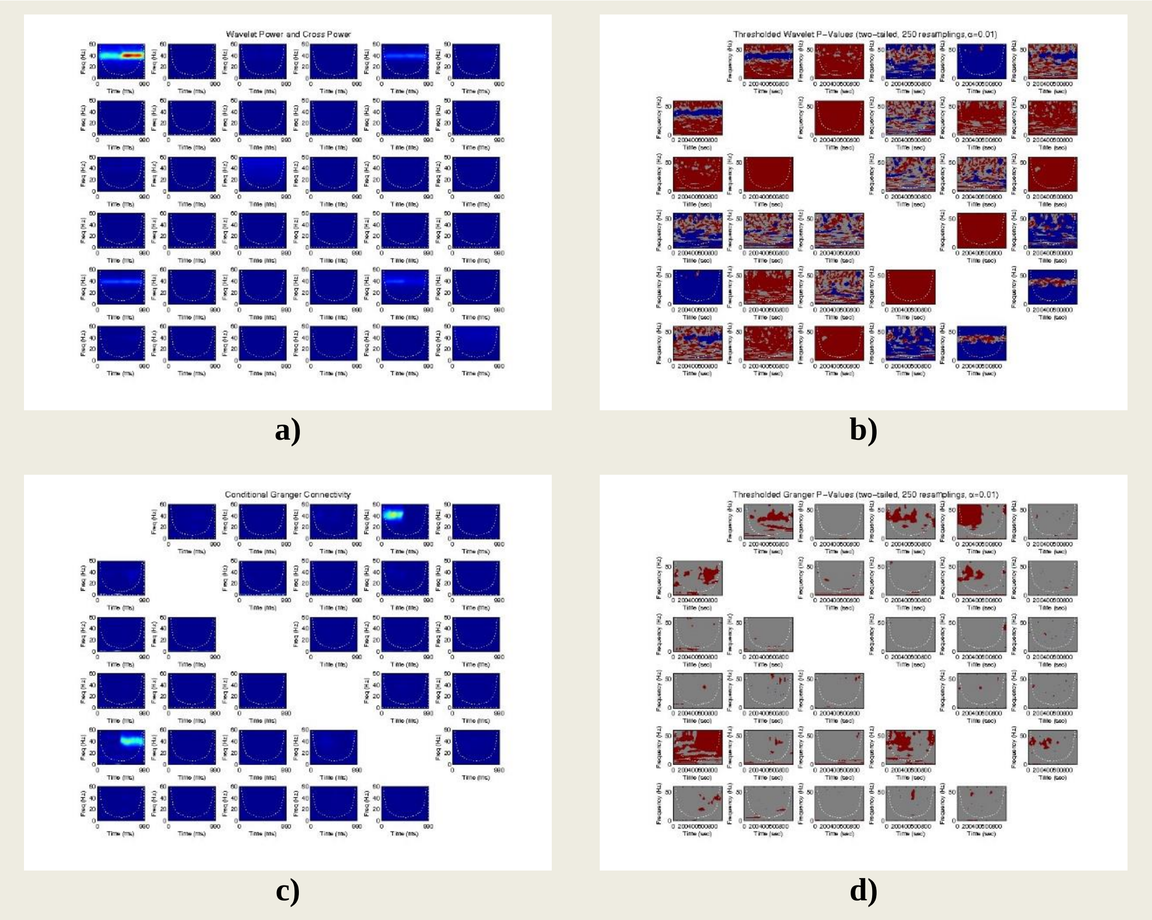
Original source localizations a), and those returned by Group-irMxNE. As opposed to the previous simulations, the locations of the original sources were not returned by Group-irMxNE.



a) b) c) d) e) f)

Figure 31: Simulation 6 Timecourses and Trials

Group-irMxNE timecourses, and individual trials for subjects ‘0‘, ‘1’, and ‘2’. This simulation utilizes three subjects with source locations and activity similar to that in Simulation 3 (Figure 20). Group-irMxNE average timecourses for each subject are shown in a), c), and e), and individual trials in b), d), and f). As opposed to the previous multi-subject simulation, individual patterns of activity are no longer clear, and this is hoped to represent Group-irMxNE beginning to combine these co-located sources into a true group solution.

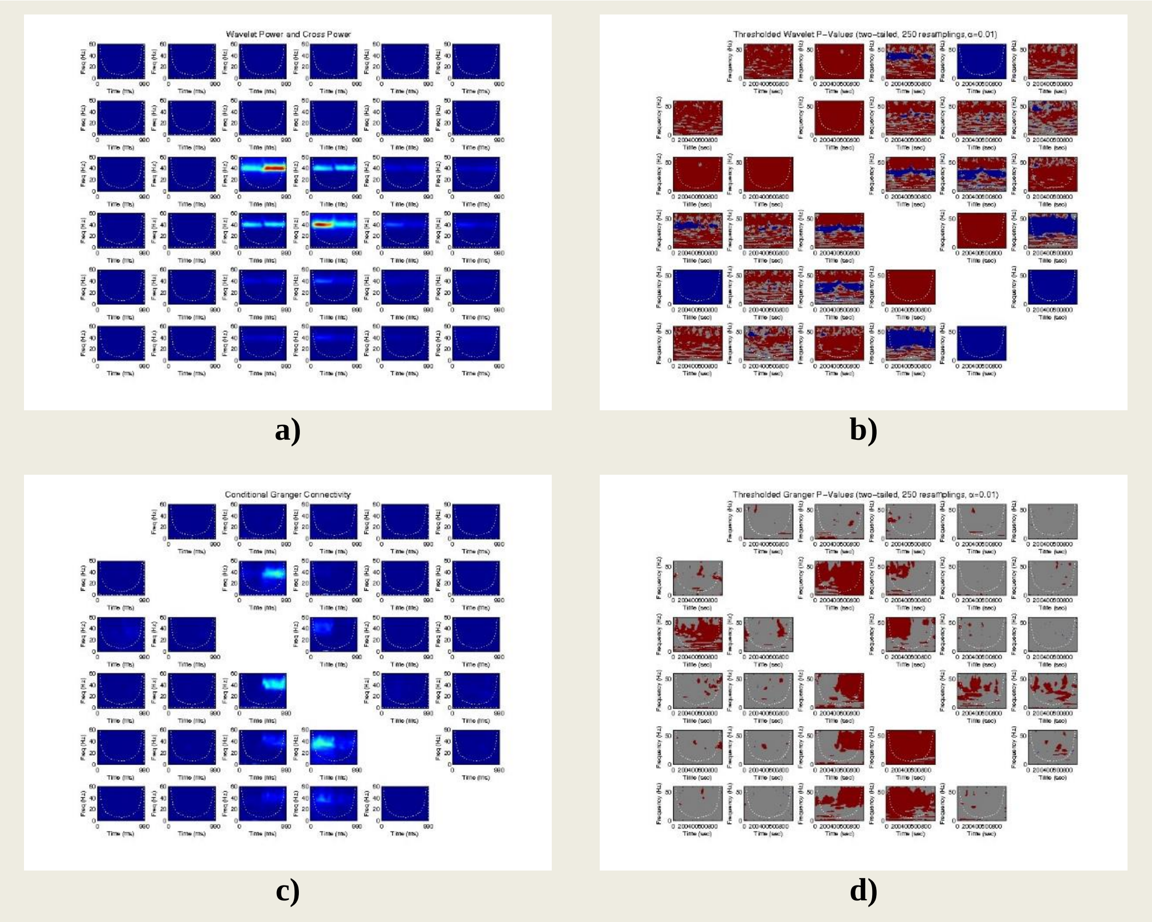


a) **b)**

c) **d)**

Figure 32: Simulation 6 Subject 1 Power and Granger Spectra

Source timecourse power and cross-power for subject '0' a), conditional Granger connectivity c), and inference b), and d). These results correspond to the individual trials output from Group-irMxNE and the multivariate regression as seen in Figure 30, b). Hints of reversing bivariate connectivity are seen in multiple sources, but are difficult to interpret.

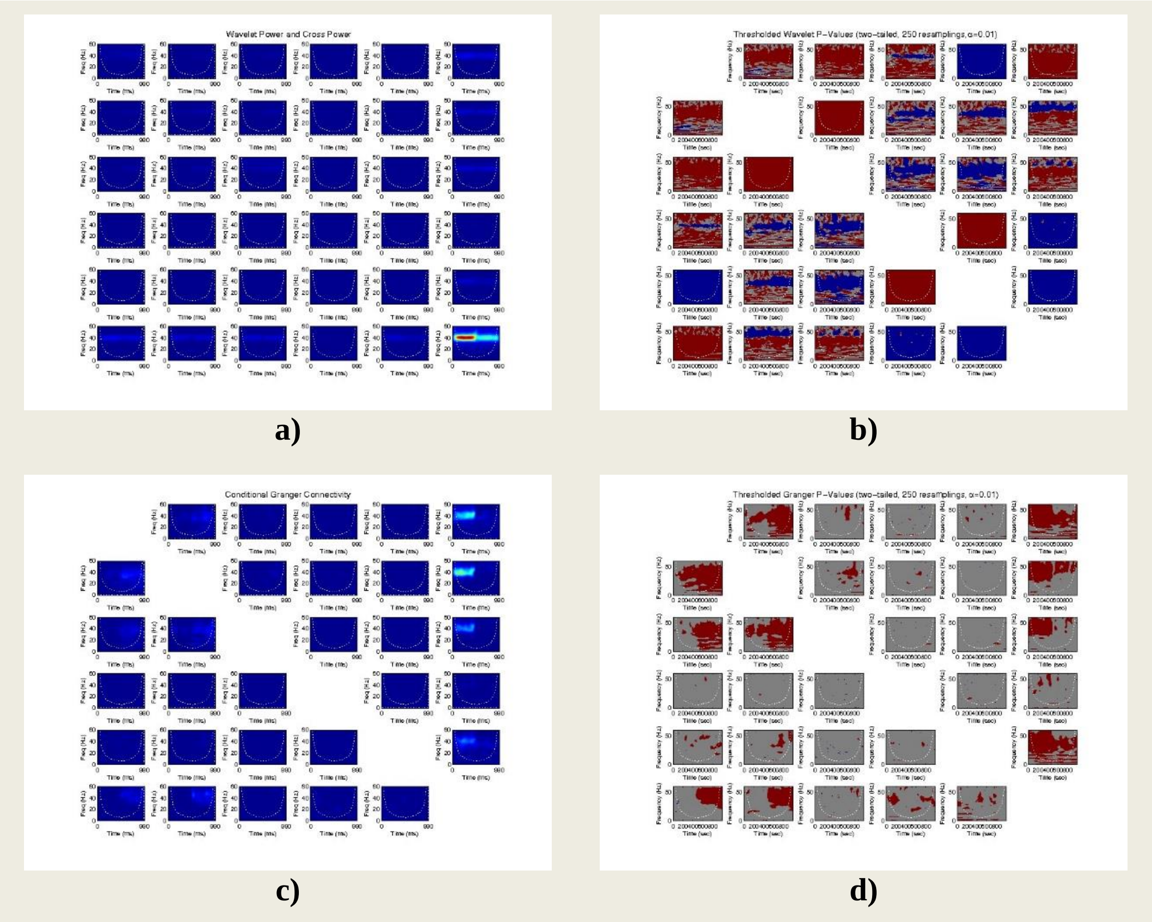


a) **b)**

c) **d)**

Figure 33: Simulation 6 Subject 2 Power and Granger Spectra

Source timecourse power and cross-power for subject ‘1’ a), conditional Granger connectivity c), and inference b), and d). These results correspond to the individual trials output from Group-irMxNE and the multivariate regression as seen in Figure 30, d). Hints of reversing bivariate connectivity are seen in multiple sources, but are difficult to interpret.

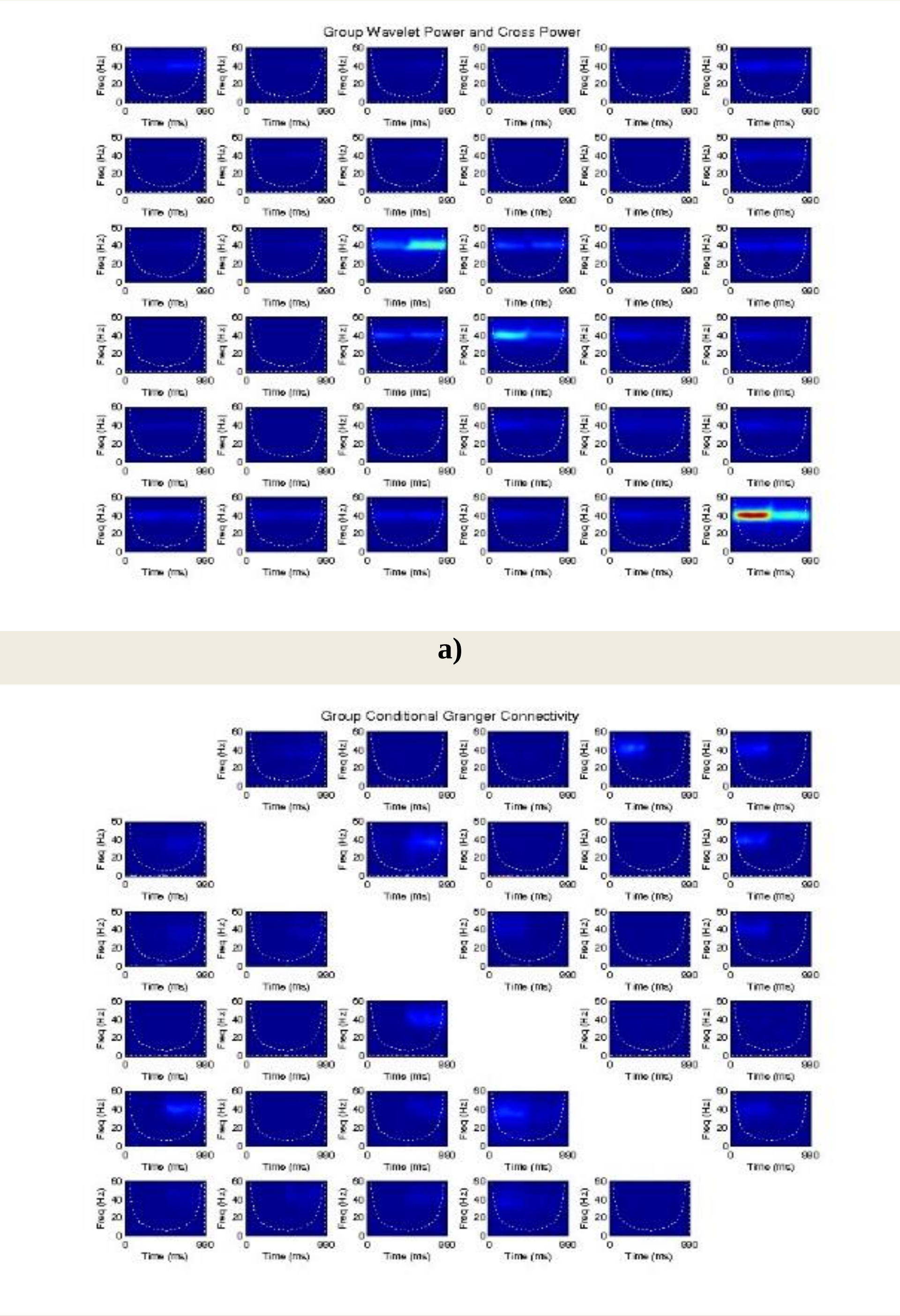


a) **b)**

c) **d)**

Figure 34: Simulation 6 Subject 3 Power and Granger Spectra

Source timecourse power and cross-power for subject ‘2’ a), conditional Granger connectivity c), and inference b), and d). These results correspond to the individual trials output from Group-irMxNE and the multivariate regression as seen in Figure 30, f). Hints of reversing bivariate connectivity are seen in multiple sources, but are difficult to interpret.



a)

b)

Figure 35: Simulation 6 Group Power and Granger Spectra

Mean source timecourse power and cross-power a), and conditional Granger connectivity b). These spectra are the simple means of those in Figures 31, 32, and 33, a) and c). In combination with the localizations seen in Figure 29, b), they represent the group model for the three subjects entered into this simulation. Overall these results show the group model to be having difficulty distinguishing between the activities of the three subjects, which is to be hoped for on the route to developing a group localization from subjects with similar source activity.

Subject	Total	Src 0	Src 1	Src 2	Src 3	Src 4	Src 5
0	**0.33** **(0.11)**	**0.38** **(0.15)**	0.00 (0.00)	**0.04** **(0.00)**	0.02 (0.00)	**0.41** **(0.16)**	**0.08** **(0.01)**
1	**0.58** **(0.34)**	**0.05** **(0.00)**	**0.13** **(0.02)**	**0.37** **(0.14)**	**0.46** **(0.21)**	**0.31** **(0.10)**	**0.16** **(0.02)**
2	**0.70** **(0.49)**	**0.13** **(0.02)**	**0.15** **(0.02)**	**0.15** **(0.02)**	**0.05** **(0.00)**	**0.13** **(0.02)**	**0.67** **(0.45)**

Table 4: Simulation 6 Goodness of Fit

Goodness of fit results for the Regression Trials Model. The group localizations are seen to account for 11% of the variance of the original evoked response of subject ‘0’, 34% for subject ‘1’, and 49% for subject ‘2’. As opposed to the results of the previous simulation, in which the model fits two subjects at similar levels (Table 5), the model here is seen to fit different subjects to differing degrees. The clean pattern of pairs of sources corresponding to unique subjects seen in the previous simulation is missing here.

Results for Facial Emotion Analysis

Results for Facial Emotion Single Subjects Analysis

The single-subject source localizations and conditional Granger spectra for the fifteen subjects with individual MRIs and head models are included in this work as appendices. See Appendix A for source localizations and Appendix B for conditional Granger spectra. Table 5 contains the goodness of fit for all twenty-three subjects. Column two gives the model fit of those fifteen subjects with individual MRIs when using their own head models. Column three is the goodness of fit for all subjects when coregistered directly onto the average brain, but when processed individually. Column four represents the goodness of fit of the group model to each subject (necessarily using the average brain).

Subject	**Individual MRI Total Model Fit**	**Average MRI Total Model Fit**	**Average MRI Total Group Model Fit**
0	**0.52 (0.27)**	**0.56 (0.31)**	**0.33 (0.11)**
1	**0.45 (0.20)**	**0.56 (0.31)**	**0.35 (0.12)**
2	**0.28 (0.08)**	**0.33 (0.11)**	**0.17 (0.03)**
3	**0.48 (0.23)**	**0.52 (0.27)**	**0.18 (0.03)**
4	**0.54 (0.29)**	**0.52 (0.27)**	**0.35 (0.12)**
5	**0.58 (0.33)**	**0.39 (0.15)**	**0.08 (0.01)**
6	**0.53 (0.28)**	**0.58 (0.34)**	**0.17 (0.03)**
7	**0.72 (0.52)**	**0.72 (0.52)**	**0.13 (0.02)**
8	**0.56 (0.31)**	**0.55 (0.30)**	**0.26 (0.07)**
9	**0.56 (0.31)**	**0.54 (0.30)**	**0.10 (0.01)**
10	**0.53 (0.29)**	**0.62 (0.38)**	**0.42 (0.18)**
11	**0.51 (0.26)**	**0.49 (0.24)**	**0.43 (0.18)**
12	**0.53 (0.28)**	**0.51 (0.26)**	**0.34 (0.12)**
13	**0.48 (0.23)**	**0.54 (0.29)**	**0.12 (0.01)**
14	**0.60 (0.36)**	**0.69 (0.47)**	**0.37 (0.14)**
15	**-**	**0.61 (0.37)**	0.02 (0.00)
16	**-**	**0.46 (0.22)**	**0.25 (0.06)**
17	**-**	**0.47 (0.22)**	**0.23 (0.05)**
18	**-**	**0.56 (0.32)**	**0.24 (0.06)**
19	**-**	**0.54 (0.29)**	**0.26 (0.07)**
20	**-**	**0.37 (0.14)**	**0.21 (0.05)**
21	**-**	**0.57 (0.33)**	**0.43 (0.19)**
22	**-**	**0.51 (0.26)**	**0.26 (0.07)**

Table 5: Facial Emotion Models Comparison of Goodness of Fit

Goodness of fit of the Regression Trials model for 23 facial emotion subjects. Column 2 shows the fit of individual-subject models for those subjects for whom individual MRIs are available. Column 3 shows the fit of individual subject models for all subjects registered directly onto the average brain, and column 4 shows the fit of the group model to all subjects. Column 2 may be compared directly to the simulations seen in Tables 1-4. There is no analog in the simulations to using brain topography that does not fit the imaging data.

Results for Facial Emotion Group Analysis

The Group-irMxNE source localizations for the model utilizing all twenty-three subjects coregistered directly to the average brain are seen in Figure 36. The average conditional Granger spectra across these subjects is found in Figure 37. Individual conditional Granger spectra for these subjects are found in Appendix C, and Figure 37 represents a simple average across these spectra. The goodness-of-fit for the complete model for each of the subjects is seen in column 4 of Table 5. Goodness of fit for each subject is broken down source-by-source in Table 6. Column 2 of Table 6 matches column 4 of Table 5, but note that Table 6 only contains raw correlation values and the reader must mentally perform the squaring operation to determine r^2 and the percentage of variance accounted for (since values are small, r^2 would be zero in its first two digits in many cases and, if both were presented as in the previous tables, Table 6 would not fit on a single page).

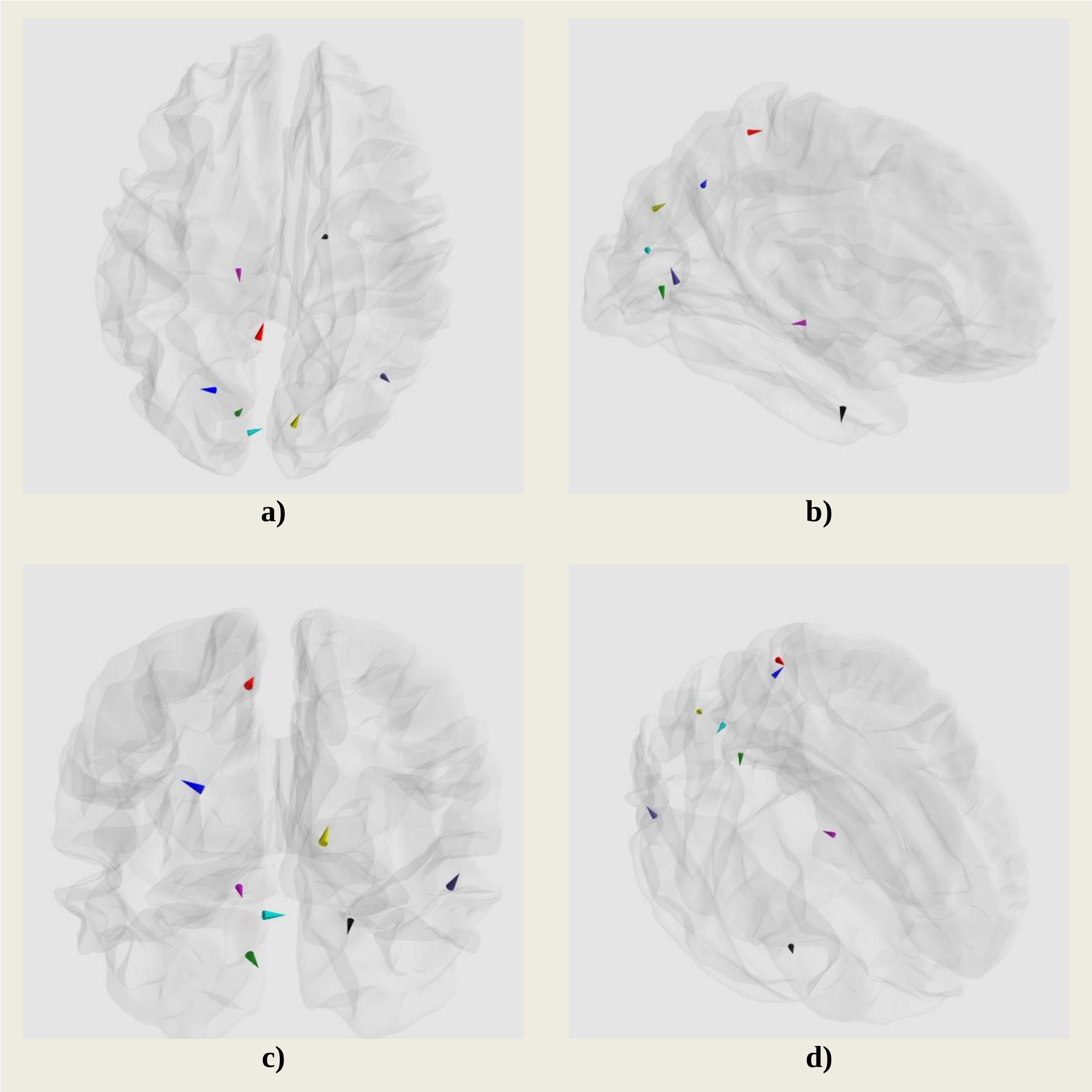


Figure 36: Facial Emotion Group Source Localizations

Group-irMxNE localizations based on 23 facial emotion subjects co-registered directly onto the average brain. These localizations may be compared to the Haxby, Hoffman, and Gobbini (Figure 7) and Adolphs models (Figure 8). Activity in the occipital region is well represented, with parietal and inferior-temporal (fusiform) sources being seen as well. The number of sources is artificially limited to eight due to computational limitations, and more frontal sources do begin to appear as the total number increases. Frames from the corresponding MNE L2-norm distributed localization (i.e. "movie") may be seen in Appendix D, which reveal that these localizations are being placed in the areas of disproportionate power.

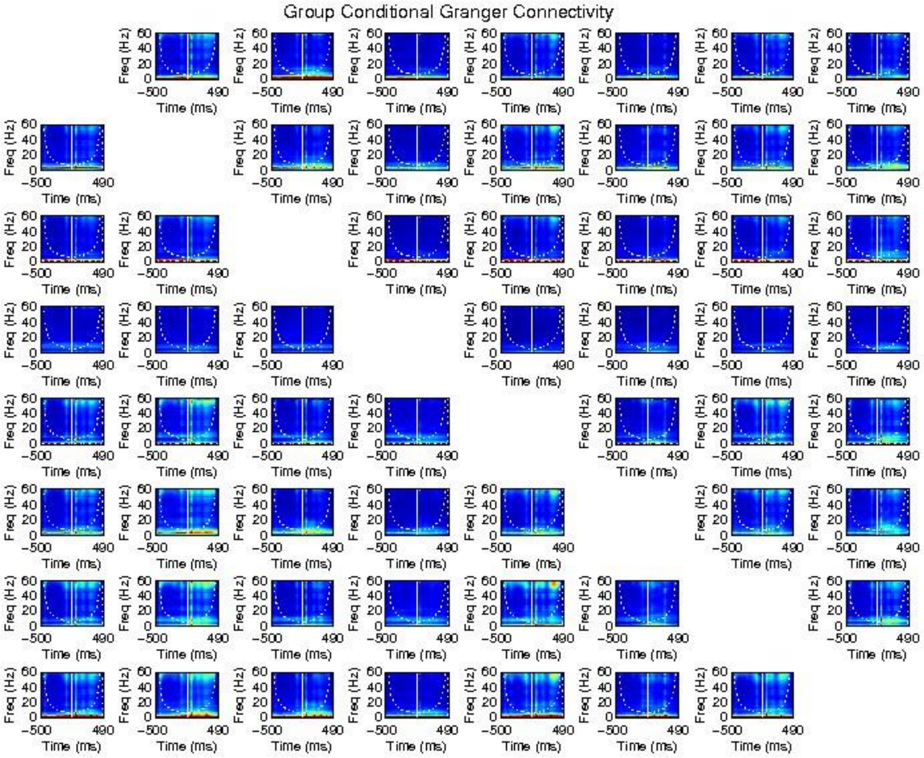


Figure 37: Facial Emotion Group Granger Spectra

Mean conditional Granger connectivity for the facial emotion group model. These spectra represent the mean of those seen in Appendix C. Mean power is not displayed because differences between subjects and noise make this difficult to interpret. Granger causality, being unrelated to power levels, may usefully be averaged. Additionally, noise in power spectra usually does not appear in connectivity results. Inference is not provided due to lack of a group mechanism. Results matching the pattern seen in the Granger spectra in Simulation 2 are seen in several places, but are difficult to interpret based on the source locations in Figure 36.

Subject	**Total**	**Src 0**	**Src 1**	**Src 2**	**Src 3**	**Src 4**	**Src 5**	**Src 6**	**Src 7**
1	**0.33**	**0.20**	**0.26**	**0.09**	**-0.04**	**0.08**	**0.22**	**0.02**	**0.24**
2	**0.35**	**0.16**	**0.19**	**0.05**	**-0.13**	**0.03**	**0.32**	**-0.05**	**0.16**
3	**0.17**	**0.11**	**0.08**	**0.05**	**0.04**	**0.08**	**0.10**	0.01	**0.09**
4	**0.18**	**0.07**	**0.23**	**0.05**	**0.15**	**0.04**	0.00	**-0.03**	**0.21**
5	**0.35**	**0.11**	**0.20**	**0.05**	**0.04**	**0.13**	**0.36**	**0.04**	**0.22**
6	**0.08**	**0.09**	-0.01	**0.08**	**0.10**	**0.05**	-0.01	**0.04**	**-0.06**
7	**0.17**	**0.11**	**0.13**	**-0.04**	**0.08**	**0.06**	**0.10**	**0.14**	**0.04**
8	**0.13**	**0.01**	**0.19**	**0.19**	0.01	**-0.04**	0.00	**-0.05**	**0.17**
9	**0.26**	**0.05**	**0.23**	**0.08**	**0.07**	**0.07**	0.02	**0.07**	**0.05**
10	**0.10**	**-0.13**	**0.03**	**0.03**	**0.07**	0.02	**0.08**	**0.04**	**0.08**
11	**0.42**	**-0.08**	-0.01	**0.09**	**0.36**	**0.14**	**0.15**	**0.12**	**0.14**
12	**0.43**	**0.18**	**0.24**	**-0.08**	-0.02	**0.07**	**0.27**	**0.14**	**0.34**
13	**0.34**	**0.09**	**0.17**	**0.09**	**0.16**	**0.08**	**0.30**	0.01	**0.19**
14	**0.12**	**-0.05**	**0.14**	-0.01	**-0.05**	**0.11**	0.02	**-0.03**	**0.05**
15	**0.37**	**0.07**	**0.09**	**0.13**	**0.10**	**0.07**	**0.44**	**-0.06**	**0.14**
16	0.02	**-0.04**	**-0.16**	**0.05**	**0.04**	**0.17**	**-0.14**	-0.00	0.02
17	**0.25**	**0.26**	**0.13**	**0.03**	-0.02	**0.04**	**0.17**	**0.04**	**0.08**
18	**0.23**	**0.04**	-0.01	**-0.13**	**0.03**	**0.07**	**0.03**	**0.17**	**0.27**
19	**0.24**	-0.01	**0.13**	**0.09**	**0.06**	**0.09**	**0.25**	**0.03**	**0.06**
20	**0.26**	**0.31**	**-0.07**	**0.11**	**-0.06**	**0.07**	**0.11**	**0.05**	**-0.11**
21	**0.21**	**0.06**	**0.10**	**0.12**	**-0.06**	**0.12**	**0.14**	**0.10**	**0.10**
22	**0.43**	**-0.03**	**0.10**	**0.11**	**0.34**	0.01	**0.35**	**0.05**	**0.13**
23	**0.26**	**0.07**	**-0.04**	**0.13**	**0.04**	**0.05**	**0.08**	**0.14**	**0.21**

Table 6: Facial Emotion Group Model Goodness of Fit

Goodness of fit results for the Regression Trials Model of the facial emotion subjects. Due to space limitations, only correlations are shown. These values are interpreted in a similar manner to those in Tables 1-4.

Chapter 4: Discussion

Discussion for Software

A significant portion of this dissertation concerns the implementation of a data processing pipeline to perform group network analysis on non-stationary MEG datasets. In order to accomplish this, a limited number of neural sources must first be identified. In order to avoid the problem of comparing disparate networks, these sources must be common between subjects, requiring the use of a group localization procedure. The localization procedure must not require the data to be stationary, must be able to operate on data containing dependent sources, and must generate the single-trial level output necessary for spectral estimation as a precursor to further processing. These requirements are met by the irMxNE sparse localizer, with extensions made to support group modeling (Group-irMxNE). A regression procedure is identified to generate single-trial output from the resulting Group-irMxNE model.

Although many measures of connectivity are employed in neuroimaging research, this work hoped from the outset to employ Granger causality. Granger causality is commonly implemented based on MVAR estimation, but the resulting use of a single estimate to represent the system over the entire analysis window precludes the analysis of dynamic systems. The assumption of static systems is implied by the definition of stationarity. This work seeks to enable the analysis of dynamic systems, motivating its avoidance of the stationarity assumption. For this reason, non-parametric Granger causality techniques, which eliminate the need for MVAR estimation, are utilized here. Increased sensitivity is seen when Granger causality measures are decomposed by frequency. The time-frequency domain enables the presentation of Granger causality results which change over time, and is utilized by this pipeline. Measures of fit to individual subjects are provided, and resampling-based inference is accomplished for individual subjects as well.

The output of this pipeline is a single set of group source localizations and a single conditional Granger causality spectra. The diagram found in Figure 9 gives an overview of the data processing pipeline. The complete pipeline is successfully tested with single and multiple subjects in Simulations 3-6, and then applied with mixed results to the MEG data collected from twenty-three facial emotion subjects.

Group Localization

Completion of this work required either a components analysis scheme or a sparse localization algorithm which is able to work with non-stationary MEG data and with dependent sources, which produces single-trial output that faithfully represents the actual source activity in the subject, and which implements a group solution. Literature searches revealed that no such functionality exists in the field of neuroimaging analysis. Fortunately, the recently-developed MxNE family of sparse localizers are designed for use with non-stationary data. Further investigation indicated that extension to a group model was possible, and a regression-based procedure for generating single-trial output suggested itself. The irMxNE solver and the modifications made in the course of this work are discussed below.

irMxNE.

The MxNE family of sparse solvers was created by the Gramfort group for the explicit purpose of operating on non-stationary data. The current work, which endeavors to create a pipeline for nonstationary analysis, wholeheartedly supports this effort. In fact, based upon literature searches undertaken at the outset of this research, it is unclear that any alternative nonstationary, sparse localization algorithm exists, and this work may not have been able to proceed without the MxNE solvers. The simulations reported herein support these localizers easily being able to recover sources in simulated data, even when multiple interdependent sources are present.

The linear model with which the MxNE localizers represent the relationship between neural sources and the MEG sensor array indicates that these solvers should leave the source activity of the subject unmodified during the localization process (Equation 5). As seen in the previous results however, the original MxNE localizer was unable to produce output timecourses with sufficient fidelity to carry the MVAR structure of the input system through to the output. Fortunately for this work, the recently released irMxNE is capable of such precision. This stage of the pipeline represented a significant hazard to the current work, because localization algorithms often give priority to source location but neglect the accuracy of timecourses, and it was not taken for granted that the necessary precision would be present.

A shortcoming of the MxNE family is the lack of support for single-trial output. Most localization algorithms are based on average evoked responses, and it is not uncommon to overlook the creation of single-trial source-level timecourses. This omission again reflects the historical focus on source localization at the expense of source activity. This oversight is particularly odd in electromagnetic imaging, where millisecond level resolution is often touted, and is increasingly critical as more complex, post-localization analyses become prevalent.

Group-irMxNE.

Group solutions are rare in neuroimaging, with group-ICA being the only known example (Calhoun V. D., Adali, Pearlson, & Pekar, 2001; Eichele, Rachakonda, Brakedal, Eikeland, & Calhoun, 2011). This is odd because neuroimaging experiments commonly use multiple subjects for the same reasons as other types of experimental work, namely increased power and external validity. The success of group-ICA in fMRI analysis indicates that sufficient between-subjects commonality exists in this mode of imaging to make group solutions useful. The limited temporal resolution of fMRI implies that commonality is only necessary in the spatial domain, however. Electromagnetic group models require such commonality to include both space and time and, as this represents the first known effort, it remains unknown whether sufficient commonality exists.

The unprecedented group extensions to irMxNE represented another significant hazard to this work. Fortunately the publications of the Gramfort group made clear that this could be accomplished based on an easily-identifiable function boundary and without modifications to the numerical minimization algorithm.

The multi-subject simulations below show that Group-irMxNE produces an output analogous to an average between subjects. Distinct source activity between subjects is reflected in the output with no effort by the solver to generalize it. Similar source activity between subjects shows hints at being combined into a single solution, but it is not clear at what point this would result in a true group solution.

Finally, it is unfortunate that, due to subject-specific whitening and dewhitening processes, individual output timecourses are created by Group-irMxNE. These individual timecourses result in individual wavelet power and Granger spectra. Although the

Granger spectra may be averaged, wavelet spectra cannot without some type of inter-subject power normalization.

irMxNE single trials regression.

irMxNE takes as input an average evoked response (i.e. sensor-level timecourses averaged over multiple experimental trials), and produces a corresponding set of timecourses for the identified sources. Subsequent processing by the present pipeline however, requires single-trial timecourses from the sources (i.e. the source activation corresponding to each of the experimental trials). The relationship between neural sources and the measurements seen by the MEG sensors is known to be linear and, as such, a good candidate for the application of regression. Multivariate multiple regression is employed such that the irMxNE source timecourses are regressed timewise onto the average sensor timecourses. The resulting regression weights may then be applied to translate sensor-level timecourses, including those of the individual experimental trials, into source activity.

Goodness of fit of the irMxNE source timecourses may be assessed by translating them back to the sensor level (using the forward operator) and comparing them to the original evoked data. In a similar procedure, the individual trial source-level timecourses may also be translated back to the sensor level. The retranslated sensor-level trials may then be reaveraged and compared to the original evoked response. This is done in order to verify the success of the multivariate multiple regression, and good fit is seen throughout the results of this work.

The success of this regression means that single-trial source timecourses are available following the application of irMxNE and Group-irMxNE. This enables follow-on processing, including the prerequisite spectral estimation to non-parametric Granger causality.

Connectivity

Many different connectivity measures are employed in neuroimaging analysis (Collura, 2008). Due to its maturity and widespread acceptance across a variety of disciplines, this work endeavored to use Granger causality unless it was somehow counter-indicated. Common belief in the neuroimaging community is that Granger causality is necessarily based upon MVAR estimation, and is therefore unsuitable for use with non-stationary, dynamic systems. Publications by Dhamala, Rangarajan, and Ding have introduced non-parametric Granger causality to the field, and awareness of this option is slowly rising (2008a; 2008b). Fortunately for this work, these authors appear to have submitted code to the Fieldtrip package implementing non-parametric, conditional Granger causality in the Fourier-frequency domain. Unfortunately their work is largely undocumented has not been fully integrated by the Fieldtrip developers. Direct support for the time-frequency domain is also missing, but this is easily rectified.

Granger Causality.

A Matlab program to accomplish spectral estimation and then submit the results for conditional Granger causality processing in Fieldtrip is implemented for use in the developed pipeline. Power and cross-power spectra are estimated by taking the wavelet transform of the individual source-level trials output from the regression procedure above, multiplying each by the requisite complex conjugate in order to generate power and cross power, and then averaging in the time-frequency domain. These spectra are then submitted to Granger estimation via the ft*_connectivity_analysis* interface.

The program is tested using the bivariate MVAR system of Dhamala, Rangarajan, and Ding. This is a simple system of order $p=2$ in which unidirectional influence passes from one variable to another. The system is made dynamic by reversing this order of influence half-way through the time window. The output of the program implemented here when applied to this system are seen in the first portion of Simulation 3, with the input timecourses seen in Figure 20, c), and the patterns of Granger and conditional Granger influence in Figure 21, c) and e). The expected spectral peak is seen at 40 Hz, and the direction of influence is seen to switch directions at the midpoint. Overall, this result represents a successful replication of that of Dhamala, Rangarajan, and Ding.

Measures of Granger influence are invariant over the amplitudes of the input timecourses. Averaging across neuroimaging subjects is perilous in most cases due to differing power levels, but this danger is thus avoided with Granger spectra. This means that the individual subjects' Granger spectra may be averaged in the final stage of the pipeline to create a group result.

Opportunities are also taken within the Granger estimation program to parallelize execution and to implement resampling-based inference.

Inference

Testing for statistical significance is the final step in most analyses, and is critical for the interpretation of results. Significance testing is undertaken within this pipeline on the wavelet and Granger spectral results. Although the end result of this pipeline is a group model, the inspection of individual-level results remains informative, including for the diagnosis of poor results. The distribution of the Granger statistic is unknown, so inference is performed using an empirical resampling process (Seth, 2010). This is implemented at the individual level, but is seen to be computationally intensive. Due to this limitation it is not implemented at the group level. For an example of individual inference applied to simulated, realistic MEG source timecourses, see Simulation 2, Figure 19, b) and d). For individual inference applied to the system of Dhamala, Rangarajan, and Ding, see Simulation 3, Figure 21, b) and d).

individual Granger causality.

Resampling-based inference is implemented based on the procedure of Seth, in which source-level single-trial timecourses are reordered between trials (2010). The source timecourses for a trial taken together represent the correct ordering. When these timecourses are shuffled between trials, and thus combined with other timecourses from differing trials, a null combination is created. Granger causality is expected to be present between sources only when the correctly grouped timecourses are used, and thus a null Granger spectra may be created. The Granger spectra based on correctly grouped trials is then compared to this null spectra for significance.

The usual criticisms of resampling methods apply in this work, namely that resampling-based inference relies exclusively on the properties of the sample at hand, and external validity may thus be questioned.

group Granger causality.

Group level inference could be implemented using resampling techniques similar to those employed at the individual level. This would entail creating a group average null distribution based on an average of null trials within the individual subjects. Although this would be very expensive computationally, group inference represents an important component of the group model, and should be implemented. It would be informative to the understanding of group models if group significance could also be tied to individual significance (i.e. are group level significant results driven by significant results in one subject, several subjects, or all subjects).

multiple testing.

Finally, although no correction for multiple testing is undertaken on the spectra in this work due to the prohibitive number of resamplings required, it is appropriate to address this topic. For a broad discussion of multiple testing in neuroimaging analyses, see the comprehensive examination of Lysne (2014).

Generally speaking, Bonferroni corrections, random-field approaches, and estimating the number of actual independent tests are techniques which apply here. Bonferroni-style corrections are frowned upon as being too conservative. For example, the dimensions of the Granger spectra produced in the simulations of this work are 1,000 time points by 61 frequencies by s^2-s spectra (where s is the number of sources), implying that 61,000 independent tests would be performed per subject per spectra. Clearly a Bonferroni correction at this level would preclude the appearance of any significant results. It is also unlikely that these tests are all independent of one another, undermining the theoretical basis for this correction.

Another approach based on random-field theory applies a threshold to the raw Granger statistic images and then tests the size of contiguous supra-threshold regions. The entire spectral image is tested at once, and the size of the largest, contiguous supra-threshold region is tested against a null distribution derived from resampling. This approach is limited by the requirements of resampling but, more importantly, also depends upon the statistic in question forming a random field with constant properties under the null distribution. This assumption is violated by both the wavelet and Granger spectra, where elongated contiguous regions of activation along the time axis are much more likely at lower frequencies.

The final approach, or correcting by an estimated number of independent tests, is the most promising in the present case. The time-frequency spectra displayed in this work contain a large amount of redundant information. As discussed previously, the discrete wavelet transform formally specifies the actual number of unique, independent spectral points required to reconstruct an original signal. This number is likely to be far less than the 61,000 points in the simulated spectra. As long as the sampling rate of the timecourses is commensurate with the maximum frequency to be examined (i.e. the timecourses are not oversampled), and the maximum frequency is realistic to the problem, this number provides a good theoretical basis for the level at which multiple testing may be controlled.

Unfortunately, any of these correction procedures place a lower bound on the number of resamplings required to achieve a null distribution with the necessary resolution to test the corrected *p*-values. This number of resamplings is beyond the reach of the time and computing resources of this analysis.

Discussion for Simulations

This work contains several simulations. The first is presented to introduce the reader to the Fourier and wavelet spectra which are referenced extensively. The second demonstrates Granger influence between realistic neuroimaging timecourses, and is presented for two reasons. The first is to estimate the appearance of influence between actual neural sources in Granger spectra so it may be recognized in the ensuing results, and the second is to address the criticism that connectivity results derived from this pipeline only apply to timecourses with MVAR structure.

Simulations 3 and 4 test the pipeline from end-to-end using simulated data for a single subject. These simulations begin by generating experimental trials for pairs of sources using the bivariate, direction switching MVAR system of Dhamala, Rangarajan, and Ding. These sources are then assigned to locations on a test cortex, and the entire simulated ‘subject’ is submitted to the analysis pipeline. As these are full-pipeline simulations, results depend upon correct performance by Group-irMxNE, the multivariate multiple regression, and Granger estimation. Simulation 3 utilizes a single subject with one pair of sources, and Simulation 4 a subject with two independent pairs.

Simulations 5 and 6 test the pipeline with multiple subjects. The ability of Group-irMxNE to generalize patterns of sources between subjects when they represent similar activity is focused on. Simulation 5 uses the two pairs of sources from Simulation 4 but, instead of assigning them to a single subject, each pair is assigned to a separate subject. Since these pairs of sources are in separate cortical regions in each subject, intuition suggests that no attempt should be made to combine them (i.e. averaging apples and oranges). Simulation 6 uses three subjects, each with a pair of sources in similar locations. It is hoped that Group-irMxNE attempts to generalize this pattern into a single group model with one pair of sources.

Simulations 3 and 4 are successful, with the pipeline returning not only the original source locations but also the source-level single trials with sufficient fidelity to reproduce the original patterns of connectivity. Since this result depends upon correct operation of numerous algorithms and procedures, this is considered a noteworthy accomplishment. Simulation 5 is also successful, in that the disparate sources of two subjects are returned in a group model nearly identical to Simulation 4 (upon which the two subjects are based). The correct pattern of influence is seen in the Granger spectra, although ambiguity increases. This is attributed to different levels of noise present in the system when two subjects instead of one are created. Results of Simulation 6 are ambiguous, with source locations produced which are spatially close but not identical to those originally input. Granger influence is seen between many different sources. This is optimistically interpreted as a rudimentary attempt by Group-irMxNE to generalize the similar pattern of sources into a single, group model, but further investigation is needed in this regard.

Discussion of Simulation 1

This simulation introduces the Fourier and wavelet spectra and contrasts the two based upon their application to the same data. The timecourses created for x_1 and x_2, with frequencies at 20, 40, 60, and 80 Hz are seen in Figure 12. Note that for x_1 the 40 and 80 Hz components have amplitude ‘1’, and the 20 and 60 Hz components ‘0.5’. In x_2 only the 40 Hz component has amplitude ‘1’, and all others ‘0.5’.

The Fourier power and cross power spectra of x_1 and x_2 are seen in Figure 13. Since these are pure sinusoidal components, each is represented by a spike at the relevant frequencies. Note that these are not the raw Fourier spectra, but power and cross power. The Fourier spectra is complex-valued, and power is generated by multiplying each point in the spectra by its complex conjugate. The power spectra for x_1 is seen in the upper left of Figure 13, and x_2 in the lower right. The magnitude of those components with original amplitude of ‘1’ have spectral power ‘1’ (i.e. 1^2). Those components with original amplitude ‘0.5’ are now seen with spectral power ‘0.25’ (0.5^2). The cross power spectra are generated by pointwise multiplication of each complex values in the raw spectra of x_1

with the complex conjugate of its counterpart in x_2, and vice-versa. As such, cross power continues to be complex-valued, and magnitude is shown in the lower left and upper right panels of Figure 13 (magnitudes of ‘0.5’ are seen where raw spectral components of amplitudes ‘1’ and ‘0.5’ are multiplied).

The wavelet spectra of the same data are seen in Figure 15 (Morlet, w_0=12). While the Fourier transform collapses over the time dimension, the wavelet transform maintains it. For this reason the panels of Figure 15 are two-dimensional. Since our signals are constant in time (i.e. stationary), horizontal wavelet peaks are seen at the expected frequencies.

The wavelet power and cross power in Figure 15 are averaged over the time dimension, and the results presented in Figure 14. These may be compared directly to the Fourier spectra in Figure 13. Note that each frequency is no longer represented by a single spike, but rather by peaks which get wider at higher frequency. The peaks overlap within the valleys between them, and the spectra does not return completely to zero.

Wavelet peak amplitude (for components of the same size) is held constant as seen here via the scalings of Liu et al. and Veleda et al. (Liu, Liang, & Weisberg, 2007; Veleda, Montagne, & Araujo, 2012)(Figure 6).

Within each wavelet spectral frame, white lines delineate the *cone of influence* (Torrence & Compo, 1998). Remember that the wavelet is an entity with a temporal extent (Figure 5), and it is useful to imagine the wavelet traversing the signals to be transformed (convolution). At each end of the signal, a portion of the wavelet falls off the end. When this occurs, the spectral value currently being produced falls within the COI and, although spectral values remain present in this region, their magnitudes should not be trusted in an absolute sense. The length of the wavelet gets longer at lower frequencies and, when frequency becomes sufficiently low and the wavelet is much longer than the signal, the COI traverses the entire spectra. The COI are maintained throughout this work simply to provide clues about the width of the wavelets at each frequency. Since Granger causality is independent of spectral scaling, they technically do not apply in this case.

Finally, the Fourier and wavelet transforms are linear operations. However, since power is the squared magnitude of the raw spectra, the transformation to power and cross power is not (this explains why single-trial output is required from irMxNE).

Wavelet power and cross power spectra, as well as Granger spectral results are encountered throughout this work. It is hoped that the explanations provided here guide the reader through their interpretation.

Discussion of Simulation 2

This simulation uses waveforms taken from actual median nerve stimulation sources. The dimensions of these waveforms are scaled on a trial-by-trial basis in amplitude and time such that the amplitude and length of the first waveform are predictive of the onset, length, and amplitude of the second.

The purposes of this simulation are twofold. First, to help identify the pattern appearing in Granger spectra when influence occurs in actual MEG data. Second, to validate the Granger causality procedures here on data other than the MVAR systems utilized in the remainder of these simulations.

The original two components used here are seen in Figure 16. A flat plot of the trials generated is seen in Figure 17, and an overplot of the trials and their averages in Figure 18.

The wavelet power and cross power, conditional Granger spectra, and inference on each are seen in Figure 19. The Granger power (upper left, lower right) of panel a) corresponds well to the original waveforms. The Granger spectra in c) shows influence between x0 and x1 (row-causing-column) in the upper right, and little activity in the lower left. This corresponds to the intended direction of influence within the simulated system. The pattern seen here is hoped to represent what will be seen in the analysis of facial emotion MEG data later in this work.

Inference on the wavelet cross spectra b) is inconclusive, but inference on the Granger spectra d) confirms the visually apparently results in c). Diagonal elements are missing from the inference panels because Granger causality is undefined within individual sources (as would appear on the diagonal), and the same code is simply applied to the power spectra.

Discussion of Simulation 3

The 300 original trials for this simulation as well as their averages are shown in the upper (x_0) and lower (x_1) frames Figure 20, c). The trials are over-plotted upon one another in blue, and their timewise average is seen in red. The visually-estimated variance of each timecourse is seen to change from the first half of the time window to the second. Referencing equations 15 and 16 it is seen that, in the first half of the time window, influence between sources is $x_0 \rightarrow x_1$, and that this influence reverses in the second half to $x_1 \rightarrow x_0$. The variance of x_1 is greater than x_0 in the first half of the time window, reflecting the positive contribution of x_0 to x_1, and this pattern reverses in the second half. The values of the actual variances are approximately $\sigma_0^2=3e\text{-}18$ and $\sigma_1^2=5.75e\text{-}18$ in the first half of the window, and the values are reversed in the second. Since the mean and variance of a wide sense stationary series are required to remain constant over time, these series clearly do not qualify.

The trials are submitted directly for Granger estimation (without passing through Group-irMxNE and the regression procedure), and their wavelet power, Granger, and conditional Granger spectra are seen in the left-hand column of Figure 21, panels a), c), and e). ***In the time-frequency spectra displayed in this work, time increases along the horizontal axis from left to right, and frequency on the vertical axis from bottom to top.*** Panel a) shows the wavelet power and cross power spectra. The frames on the diagonal of this panel show the power of the timecourses of the individual sources. The off-diagonal frames show the cross power, and the entire panel is symmetric about the diagonal. (The magnitude of the cross power spectra $x_0 \rightarrow x_1$ and $x_1 \rightarrow x_0$ are identical - separating them into an additive, directional combination is the purpose of the Wilson algorithm.) Note that the power spectra for x_0 and x_1, on the diagonal of this frame, reflect the variance changes seen in panel c) of Figure 20. Panel c) of Figure 21 shows the Granger causality spectra of the simulated system. ***The diagonal elements are omitted in Granger spectra, because these are only defined between sources.*** The upper right frame of panel c) shows Granger causality $x_0 \rightarrow x_1$, and the lower left frame shows $x_1 \rightarrow x_0$. As defined in the original MVAR system, $x_0 \rightarrow x_1$ is seen in the first half of the time window, and $x_1 \rightarrow x_0$ in the second. The conditional Granger causality spectra is shown in panel e) of Figure 21, and is seen to match that of the non-conditional in panel c). This is as expected in a bivariate system with no opportunity for prima facie errors. The peak of both the wavelet and Granger spectra is seen at approximately the expected 40 Hz. These results replicate those of Dhamala, Rangarajan, and Ding (2008a; 2008b), indicating that the non-parametric Granger causality program is functioning correctly.

The right-hand column of Figure 21 shows inference results for the spectra in the left hand column. Inference is performed using the resampling procedure described previously. 250 reorderings are employed, and significance is based on $\alpha=0.01$ in a two-tailed test. Spectral locations which are significantly greater than the null are shown in red, and those which are less in blue. Inference on the wavelet spectra are included for completeness in panel b) but, since this work focuses on Granger causality, these are not interpreted. Panels d) and f) show significant Granger and conditional Granger causality occurs according to the expected pattern for this system.

The convention used throughout this work for referencing individual spectra in figures such as Figure 21, panel a) is to count rows and columns from zero starting at the upper left corner. Note that this is matrix-style indexing (row, column), and not Cartesian (x, y) starting at the lower left. For example, within this panel, the upper left spectra is the power for source x_0, and the lower right is the power for source x_1. The lower left frame is the cross power between x_1 and x_0, and the upper right the cross power between x_0 and x_1. Although the cross power is diagonally symmetric, the Granger spectra are not. Within Granger panels, the convention is 'row-causes-column'. Therefore, the lower left frame of Figure 21, panel a) shows the Granger causality spectra for $x_1 \rightarrow x_0$, and the upper right $x_0 \rightarrow x_1$.

These simulated trials are assigned to the auditory cortical sources shown in Figure 20, panel a) and passed to the pipeline of this work (x_0: left auditory, blue, x_1: right auditory, green). After several iterations required to set the regularization parameter (λ) such that Group-irMxNE returns two sources, it is seen that the two input locations are replicated. (The regularization parameter adjusts the weight of the penalty which the cost function assesses for each additional source, and therefore controls the number of sources present in the output. The necessity of setting this parameter based on prior knowledge of the number of sources, versus determining it from the data, represents a distinct weakness of this pipeline.) The timecourses output from Group-irMxNE associated with these sources are shown in solid lines in panel b) of Figure 21. ***(Within***

Figure 20, and in similar figures throughout these results, the color of the sources in panel a) can be matched to timecourses in panel b).)

The multivariate multiple regression procedure is then invoked to model the relationship between the sensor-level evoked response input to Group-irMxNE and the output timecourses. The individual sensor-level trials are then propagated through this model, resulting in individual trials at the source level. These individual trials are shown in panel d) of Figure 20, and are seen to closely match the original trials in panel c). The averages of the x_0 and x_1 trials output from the regression procedure are seen in red in the upper and lower frames of panel d), and are also plotted in dotted lines in panel b). The dotted lines are largely obscured by the original solid lines, indicating that the regression model is a good fit. ***(The index of the timecourses in panel b), counting from zero starting at the top of the legend, matches those in the rows and columns of the spectra.)*** Finally, the output single trials are submitted to wavelet and Granger spectral estimation. The output is seen in Figure 22, and which may be compared directly to Figure 21. This comparison is extremely favorable, and shows that the connectivity properties of the original system are recovered by the non-parametric Granger causality pipeline.

The overall fit of the irMxNE model to the original sources and trials is addressed in Table 1. For the single subject in Simulation 3, the row for subject ‘0’ gives $r(r^2)$ for various parts of the model. The column ‘irMxNE Total’, where r^2=0.71, indicates that the two source timecourses output from Group-irMxNE and translated back to the sensor-space via the forward operator, account for 71% of the original variance. Interestingly, the output timecourses for sources ‘0’ and ‘1’ (left and right auditory cortices), account for uneven portions of this variance, at 7% and 65% respectively. The reason for this is unknown, although uneven patterns of variance favoring both sides were seen in multiple iterations of this simulation. Model fit is again assessed using the average of the individual trials propagated via the multivariate multiple regression, and is seen to match that of Group-irMxNE within two digits of precision (‘Trials Model’). This indicates that little or no variance is lost in the course of this regression. Accounting for 71% of the original variance indicates that these models are a good fit to the original data.

Generally speaking the fits for this simulation are good, although they are achieved with a signal-to-noise ratio that is unrealistic for measured MEG data. It was noted that model fits approached 100% variance-accounted-for as the signal-to-noise ratio of the trials was decreased even further, but is also expected that this level of fit is not achievable with actual MEG data.

The values of several parameters required as input to irMxNE were determined from experience with this simulation as follows: *maxit=150*, *n_mxne_iter=15*, *active_set_size=100*, *tol=1e-14*, and *weights_min=0.01*. These parameters represent the maximum number of iterations, sub-iterations, number of simultaneously active sources, cost function threshold below which to end looping early, and the minimum value of the L2-norm weights below which sources are automatically considered inactive. As opposed to the regularization parameter, these values are only seen to impact the solution when one or more of them is set too conservatively. These parameters do, however, collectively impact the total running time of the solver when set too liberally. A strategy is therefore adopted of beginning with conservative values and successively loosening them until the output solution stops changing. These values are used throughout the simulations, and only *weights_min* and *active_set_size* are changed for use on the facial emotion data.

The overall results of this simulation represent an end-to-end validation of the non-parametric, Granger causality pipeline. That the two original source locations are returned by Group-irMxNE indicates not only that irMxNE itself functions exceptionally well, but that the group extension does not disrupt its basic functionality. Even more impressively, the Granger and conditional Granger causality spectra of the original bivariate non-stationary MVAR system are also reproduced in the output (compare Figures 21 and 22). This indicates several things. First, not only did Group-irMxNE return the original source locations, but also the timecourses of these sources with excellent precision. Second, the multivariate multiple regression procedure is able to accurately model the relationship between the sensor-space evoked data and the source-space average timecourses. This model is then sufficient to translate the sensor-space trials to the source-level. The conditional Granger causality function, upon receiving a collection of source trials nearly identical to those originally simulated, produces very similar Granger and conditional Granger spectra. Finally, that the output Granger spectra reflect the changing direction of influence in the original system indicates that all of the components of the pipeline do indeed function correctly with non-stationary data. (The author has constructed a number of analysis pipelines which, although theoretically supported, have produced nothing but noise at the output. Seeing these results for the first time was indeed a relief.)

Discussion of Simulation 4

This simulation utilizes two, interacting pairs of sources, one in the left and right auditory cortices and one in the primary visual areas. The results seen here are interpreted in a similar manner to Simulation 3. The original source locations are returned by Group-irMxNE, and are seen in Figure 23, c). The input and output trials are seen in panels c) and d) respectively, and the timecourses output from Group-irMxNE as well as the average of the regressed single trials in panel b). The auditory sources are indexed as ‘0’ and ‘2’, and the visual ‘1’ and ‘3’. Figures 24 and 25 represent the wavelet and Granger spectra of the raw, simulated input data and the output from the connectivity pipeline. Within the spectra, the pattern seen in Simulation 3 is repeated between sources ‘0’ and ‘2’, and ‘1’ and ‘3’. Frames showing activity between other pairs of sources are empty except for the appearance of sporadic noise. This pattern is reflected in the inference results as well. As with Simulation 3, excellent correspondence is seen between the spectra of the input and output trials, indicating that Group-irMxNE has again produced a good solution. The Granger and conditional Granger results also continue to agree.

Goodness of fit for the stimulated results are shown in Table 2. Since the sensor-level evoked data reconstructed from the average of the regressed single trials (‘Trials Model’) subsumes that model generated directly from the Group-irMxNE output (‘irMxNE Model’), only the former is shown here. (Although the correspondence between the two models is monitored throughout this work, in order to monitor the performance of the regression procedure.) The total original variance accounted for by the single subject is 67%. The individual sources account for 13%, 20%, 21%, and 13%, respectively. That the total variance accounted for is nearly the same as Simulation 3 reflects the same level of noise being added in each case.

The auditory and visual cortices are chosen in this simulation because they represent distinct brain function and regions. When distinct patterns of sources are submitted to Group-irMxNE, it is desirable that the localizer respect these as separate locations and make no attempt to combine them. This is the result seen here. The regularization parameter (λ) is again tuned over repetitions of this simulation such that four sources are produced. However, it is noted that when this parameter is set too conservatively (i.e. fewer sources), Group-irMxNE returns some subset of the original sources, versus attempting to find a compromise among them. This is surprising because, if λ were set to produce two sources, the best fit may be represented by one source in each hemisphere at some midpoint between the auditory and visual areas. (Indeed, with λ set to produce one source, the localization is to one of the original sources.) It is suspected that this extreme behavior is a result of the independent MVAR systems used here. While it is tempting to focus on source localizations, it is important to remember that irMxNE produces localizations and timecourses as part of a single solution. Although it is intuitive that, when forced to produce a solution with a restricted number of sources, irMxNE would attempt to find some type of spatial accommodation, the fact that the activity of these pairs of sources is completely independent of one another may render this impossible. In other words, while sources at some midpoint between either the two hemispheres or the auditory and visual cortices represent a spatial compromise, that the timecourses of these sources have nothing in common may preclude this. Attempting this simulation using data such as that in Simulation 2, where there is some commonality between the activities of the sources, may produce different results.

When λ is set too liberally, allowing greater than four sources into the localization output, the additional sources are placed in proximity to the correct originals. This is taken to be a result of the noise introduced into the system, which is assumed to cause ambiguity in the localizations (similarly to the effect of noise upon the point spread function of an optical system; however, this phenomenon may also be similar to the splitting of sources seen in independent components analysis.)

Discussion of Simulation 5

This simulation utilizes the same configuration of sources and systems as seen in Simulation 4 but, in this case, each pair of sources is assigned to a separate subject. The pair of auditory sources are assigned to subject ‘0’ and the visual to subject ‘1’. The data itself is identical to the previous simulation.

Employed here in a group fashion for the first time, Group-irMxNE returns the same set of original sources as seen in Simulation 4 (Figure 23, panel a)). The input trials match Simulation 4 as seen in Figure 23, panel c) but are assigned to different subjects (subject 0: sources ‘0’, ‘2’, subject 1: sources ‘1’, ‘3’). In this case there are two separate sets of simulated trials, each belonging to a different subject, versus the data being combined as in Simulation 4. The multivariate multiple regression procedure is used twice, to regress the Group-irMxNE output timecourses (Figure 26, panels a) and c)) onto the sensor-level average evoked response of each subject (remember that the Group-irMxNE timecourses only differ between subjects according to the individual whitening and dewhitening operations, and see the data flow diagram in Figure 9), and produces two solutions. The sensor-level trials for each subject are then generated using these models, and are seen in Figure 26, panels b) and d). Since the group model contains four sources derived from the combined subjects, but each subject only contains two original sources, these panels each show two active sources and two idle ones. For example, panel b) of Figure 26 shows sources ‘0’ and ‘2’ to be active. These correspond to the auditory system assigned to the first subject. The pattern in panel d) reflects that of the second subject. Note that the idle source timecourses contain a constant level of random activation with variance slightly less than that of the active sources. This is a result of the noise added to the system (without noise these traces go completely to zero, causing spectral estimation to fail with the resulting non-invertible matrices).

Wavelet power and conditional Granger spectra for both subjects are seen in Figures 27 and 28. (Since the standard Granger causality spectra is seen in previous simulations to very closely match that of conditional Granger causality, it is not shown here.) The conditional Granger spectra and its corresponding inference for subject ‘0’ is seen in panels c) and d) of Figure 27, and continues to display the pattern established for the system first seen in Simulation 3 (between sources ‘0’ and ‘2’). The Granger spectra of subject ‘2’ seen in Figure 28 shows a similar pattern of influence between sources ‘1’ and ‘3’. Large regions of noise are now seen in the Granger spectra, with the conditional Granger spectra of subject ‘0’ showing influence of source ‘3’ upon sources ‘0’ and ‘2’ (Figure 27, panels c) and d)). Similar noise in the conditional Granger spectra is seen for subject ‘2’, with source ‘2’ appearing to influence source ‘1’ (Figure 28, panels c) and d)).

Two related factors change in Simulation 5 which may cause greater noise than seen in Simulation 4. First, two subjects are input instead of one and second, the simulated sensor-level trials are split between subjects, causing greater effective noise levels. The data generated for Simulation 4 consists of 300 trials, with each trial including contributions from four sources. 3 dB of noise is added to each trial. In Simulation 5, 600 trials are generated, with 300 for each subject. The trials for each subject contain the contributions of the pair of sources assigned to that subject, but not for pair belonging to the other subject. The same 3 dB of noise is added to all 600 trials, of which now only half contain the contributions of each pair of sources. On this basis the total noise in the system is double that of Simulation 4. This is reinforced by the observation (not shown) that the noise seen in the conditional Granger spectra of both subjects is reduced as the noise power is decreased.

Goodness of fit is seen to be reduced in Simulation 5 versus Simulation 4, likely for the reasons of differing noise discussed previously. Table 3 shows the percentage of variance accounted for by the group model with regard to subject ‘0’s original evoked response to be 54%, and subject ‘1’s 58%. Contributions of the individual sources to each subject are shown as before and, as expected, only two sources contribute to each subject. For subject ‘0’, sources ‘0’ and ‘2’ contribute 25% and 41% variance, respectively, with sources ‘1’ and ‘3’ contributing 0% apiece. For subject ‘1’, the pattern is reversed, with sources ‘1’ and ‘3’ contributing 48% and 20%, and sources ‘0’ and ‘2’ 0% each. This pattern reflects the pairs of sources assigned to each subject, with sources originating in the opposite subject being inactive. (The four group-localized sources are being energized by the individual subjects’ data and, since they are spatially distinct, only the two sources present in each subject are active. See Figure 9.) The utility these individual-subject, individual-source contributions in assessing which sources are most active in which subjects is seen here.

Group average wavelet power and conditional Granger spectra are seen in Figure 29. Since group inference is not provided by the pipeline, no significance spectra are shown. This figure may be compared directly to Figures 24 and 25, panels a) and e) in Simulation 4, where the same pattern of sources and influence are present in a single subject. Both wavelet power and conditional Granger spectra compare very favorably, showing that the group model of two subjects closely matches the model of a single subject with the same sources. Note that the measures of Granger influence seen in the group model (Figure 29, b)) show the effects of averaging when compared to those of the single-subject model (Figures 24 and 25, e)). The amplitude of the group Granger spectra is reduced by being averaged across two subjects, when the values being averaged exist in only one subject. For example, influence from source ‘2’ to source ‘0’ is only present in subject ‘0’, and the Granger values for this subject match those of Simulation 4. However, these values are then averaged with the spectra for subject ‘2’, for whom this panel is empty. This is the expected behavior for a group model.

Overall this simulation validates the performance of Group-irMxNE on multiple subjects. Since source activity within the subjects is spatially distinct, there is little interference between the two patterns, and the output closely matches that seen in Simulation 4.

Discussion of Simulation 6

Like Simulation 5, this simulation utilizes multiple subjects. Unlike the previous simulation however, where two subjects with distinct pairs of sources are created, Simulation 6 utilizes three subjects, each with a pair of auditory cortical sources as in Simulation 3. As opposed to distinct sources which intuition suggests should remain separate in a group model, the three pairs here are intended to be similar. In this case all three subjects are engaged in similar activity, which could be generalized into a single system in the group model.

The original source localizations are not reproduced by Group-irMxNE, and Figure 30, panels a) and b) show the original and returned locations. The Group-irMxNE source locations are seen to be similar, but not identical. The discussion of noise in Simulation 5 applies here as well, but with the addition of a third subject. For this reason the original signals may be even further reduced, leading to this result.

In Simulation 5 the original sources are seen to be correctly, unambiguously separated by subject. That is not the case here, and Figure 31 shows the Group-irMxNE output timecourses and the single trials. A unique pair of sources appears to be assigned to subject ‘1’ (Figure 31 panels b) and c)), but for subjects ‘0’ and ‘2’ the pattern is ambiguous. The wavelet power and conditional Granger spectra for the three subjects and seen in Figures 32, 33, and 34 show the same results. Here the pattern for subject ‘0’ in Figure 32 is the most clear, with sources ‘0’ and ‘4’ somewhat replicating the expected pattern. Subject ‘1’, seen in Figure 33, also replicates the pattern, but with multiple sources participating. In particular, causality is seen such that ‘1’→’3’ and ‘2’→’3’ in the first half of the time window, and ‘3’→’2’, ‘4’→’2’, and ‘5’→’2’ in the second (most easily seen in panel d)). The final subject in Figure 34 shows similar patterns with ‘0’→’5’, ‘1’→’5’, and ‘2’→’5’ in the first half of the window and ‘5’→’0’ and ‘5’→’1’ in the second. For all subjects this pattern of causality appears more faintly between additional pairs of sources not listed here. In Simulation 5, with distinct pairs of sources, this would be a disappointing result. With these collocated pairs however, these results are encouraging, indicating that the solver is attempting to generalize the activity of all three subjects in some way. Despite being in similar locations however, the three sources in each hemisphere do not have similar orientations, as is reflected by the arrows in both panels of Figure 30. Due to this, the magnetic fields generated by activity in these sources are not expected to be completely similar, and the sources are therefore not interchangeable in the way that could lead to their being combined in a group model.

Goodness of fit for Simulation 6, as seen in Table 4, reveals that the model does not fit the subjects evenly. The group model accounts for 11% of the original variance for subject ‘0’, 34% for subject ‘1’, and 49% for subject ‘2’. This is disappointing given that the simulated sources for all three subjects have identical power and noise. The sources represented in green and blue in the left hemisphere of Figure 30, b) have similar orientation, and perhaps two subjects are well matched in this regard, but the third is the ‘odd man out’. The individual source results for subject ‘0’ show that sources ‘0’ and ‘4’ have the only notable activity, and these may represent the best fit to the two original sources for this subject. (While 0 and 3 were the original source locations corresponding to this subject, Group-irMxNE identifies sources in no particular order.) Subjects 1 and 2 appear to have a greater degree of overlapping source activity in sources 1, 2, 4 and 5, further supporting this hypothesis. As suggested previously, it would be interesting to repeat this simulation with more realistic source activity such as that generated in Simulation 2.

The group average power and conditional Granger spectra for this simulation are seen in Figure 35.

Generally speaking, while this simulations reveals Group-irMxNE to show some hints of combining the sources of subjects' with similar activity, this effect needs to be better understood.

Summary and Conclusions for Simulations

These simulations generally support the correct functionality of the group non-stationary connectivity pipeline constructed here. Multiple simulations of the known MVAR system of Dhamala, Rangarajan, and Ding show that not only are these authors' Granger connectivity results replicated, but that these results may be recovered, along with correct localizations, from simulated subjects' data. The pipeline functions correctly for single-subjects with multiple sources, and for two subjects with distinct sources. When presented with three subjects with similar, collected sources, Group-irMxNE neither recovers the original sources and pattern of connectivity, nor generalizes them into a single system. Further work is needed in this regard, and a principled method for determining the number of sources to represent in the output also should be developed.

The expected pattern of connectivity is also returned from Granger analysis of realistic MEG waveforms and, although not tested with the full pipeline, it is expected that success would be seen here as well.

Discussion of Facial Emotion Analysis

Although seemingly counterintuitive, good practice dictates that individual-subjects analysis be a precursor to group modeling. Fortunately, all of the preprocessing steps are shared, and individual submission to the pipeline is the only difference. Subjects showing an individual pattern of sources which is not at least generally correct for the facial emotion protocol, or which result in excessively noisy Granger spectra, should be considered for additional data cleaning or exclusion (none of the facial emotion subjects were). Following group modeling, it may be desirable to revisit the individual models and to tune their regularization parameters to generate the same number of individual sources as does the group model. This facilitates a direct comparison of the goodness of fit of the individual models and group model to the individual subjects. Such a comparison is useful when diagnosing uneven fit in the group model, where it is useful to know if the individual model fits poorly as well.

This work is confronted with the additional challenge that the morphing capability in MNE, needed to transform individual subjects' head models into a common space, is unavailable. For this reason it is desirable to estimate the impact of co-registering subjects' MEG data directly onto the average brain, and this is done by analyzing individual subjects using both their own MRI-derived head shapes (where available), and the average brain. These results are discussed along with the group facial emotion model below.

The pipeline developed here is experimental, particularly with regard to group modeling of MEG data. The models of the fifteen facial emotion subjects with individual MRIs and head models provide an upper bound on what may be accomplished with this pipeline, and there are presented in the results of this work for comparison to the group model.

Discussion of Facial Emotion Single Subjects Analysis

Localizations from the single-subjects analysis seen in Appendix A show a uniform pattern of sources in the occipital and surrounding regions. Occasional sources are seen wrapping forwards into temporal and, even less frequently, frontal areas. The MxNE cost function (Equation 6) makes clear that solver attempts to generate sources and timecourses in such a way as to best estimate the given evoked response. This implies that preference is given to locations representing the greatest neural activity over the window being modeled. Frames from the MNE L2-norm distributed localization (i.e. "movie") are seen in appendix C, and indicate that the greatest power in the combined evoked response occurs in the time window of 100-200ms post-stimulus and centers around the occipital, striate, posterior temporal, and inferior parietal regions. The sensor-level average evoked response seen in Figure 2 also confirms this observation. On this basis it is not surprising that source localizations are focused in and around the occipital cortex. The eight-source threshold applied to all of the localizations in this work represents an arbitrary limit beyond with interpretation becomes difficult. The question is suggested that additional, frontal sources may be revealed with a higher threshold. Although figures are not presented here, experience with the solver indicates that, as the number of sources is allowed to increase, frontal sources are indeed seen more frequently, but the number of near-occipital sources increases as well.

The early stages of the Adolphs model are well represented in the single-subjects analysis. Although not specifically noted by Adolphs, visual processing in occipital regions such as the calcarine fissure must be the first stages of any visual-stimulus model, and multiple sources are seen in this region in most subjects. Cortical processing in the Adolphs model formally begins with the striate, which is also well-represented, and early processing in the inferior and superior temporal areas is seen in many subjects. Disappointingly, more frontal localizations which would presumably correspond to the later stages of activation are only seen sporadically and too infrequently to discern a pattern.

It would be naïve to expect a one-to-one correspondence between the regions noted in the Adolphs model and sources output from irMxNE. On the other hand, there is clearly a disconnect between theoretical, information-centric models of neural processing and power-centric modeling of evoked responses as represented by most localization algorithms. Efforts on both fronts are suggested, where authors of information-processing models are encouraged to be more comprehensive in their works (e.g. Adolphs' exclusion of visual processing, which accounts for a large amount of power), and developers of source localization algorithms to be cognizant of the need to be more sophisticated than the simple modeling of power. Along the later lines, the current results indicate that some type of power normalization over time may be usefully performed on the evoked responses prior to source localization with irMxNE.

Conditional Granger spectra resulting from the single-subjects analysis are seen in Appendix B. These spectra are very encouraging in the general sense that instances of the expected pattern of connectivity (Simulation 2, Figure 19) are seen at least once for nearly every subject. The appearance of this pattern suggests that actual instances of influence between sources are being successfully identified by the pipeline. Inference on the spectra seen in Appendix B is unfortunately omitted due to time and computational constraints. If the visually apparent indications of connectivity seen here were supported by statistical significance, this conclusion would be further upheld. Detailed interpretation of patterns of connectivity are not undertaken due to both the lack of a theoretical framework supporting the found locations of sources, and the idiographic nature of the results (i.e. this is exactly the type of network comparison problem that the group model seeks to avoid).

Goodness of fit for the single-subjects models seen in Appendices A and B are found in column 2 of Table 5. For the worst-fit subject, 8% of the variance in the original evoked response is accounted for (subject ‘2’), rising to 52% in the best fit case (subject ‘7’). Values in the range of 30% are common. Assessed as correlations, the fit of most subjects falls in the moderate range (0.5), with the occasional high (0.7) and low (0.3) value being seen. Given that these values are based on models with artificially-limited numbers of sources, these results are taken to be encouraging.

Generally speaking, the good fit and reasonable source localizations seen in the single-subjects analysis are a positive development. Assuming that the conditional Granger spectral results are supported by inference, further development of this pipeline for use with individual subjects seems to be a productive endeavor. Work is still needed however, to include power normalization of evoked responses over time, identification of a principled means to choose the number of sources, and validation on an established neural model of connectivity such as the median nerve stimulation protocol.

Discussion of Facial Emotion Group Analysis

Not surprisingly, the sources identified by Group-irMxNE across the twenty-three facial emotion subjects mirror the pattern seen in the individual results (Figure 34). Visual processing is represented with several occipital sources near the calcarine fissure (left: light blue, dark green, right: yellow). A source in the right striate cortex indicates the starting point of processing in the Adolphs model (dark blue), and sources appear near the inferior surface of the temporal lobes (left: purple, right: black). Two left parietal sources are seen as well (blue, red). These eight sources are organized in Table 7. It is disappointing that the right inferior fusiform gyrus, which is seen to be a center of significant activation in a single subject in Figure 4, and also in the group snapshots of Appendix D in the time window 150-200ms, is not represented in the source localizations.

The individual conditional Granger spectra output by the group model are found in Appendix C, and several subjects are seen to display hints of the expected pattern of connectivity found in Simulation 2 (Figure 19). Other subjects are seen to be notably noisy. The appearance of noise to this degree is surprising since Granger influence is a complex process and one unexpected to occur randomly. No attempt is made to interpret the individual Granger spectra other than to note that several might be considered for exclusion from the average.

The group average conditional Granger spectra is seen in Figure 35. The spectra seen here represents a simple average of those in Appendix C. The fortunate property that measures of Granger influence are independent of either the absolute or relative power levels of the power spectra used to create them renders such an averaging procedure reasonable between non-normalized subjects. Hints at the expected pattern of connectivity are seen in the group average spectra as well. The subspectra at position ‘5, 1’ is the best example of this, but also implies that the right superior occipital source (yellow) is influencing one of the left primary visual sources (green). Even if this causation didn’t cross hemispheres, it would be occurring opposite of the expected direction. Another example, seen at position ‘1, 4’, implies that one of the left primary visual sources (green) is influencing the left inferior medial temporal source (magenta). Although the temporal resolution seen in the Granger spectra is limited (partially due to the choice of wavelet parameter $w_0=12$), influence in subspectra ‘1,4 ‘ appears to occur relatively late in the response, rendering the timing correct for a forward activation of the fusiform region (although direct primary visual to fusiform activation is not predicted by the Adolphs model). Had this influence occurred very early, and in the opposite direction, it would suggest that source ‘4’ is actually the left amygdala influencing visual processing in early stages of the Adolphs model. Along with seeking to improve the temporal resolution of the wavelet spectra, reanalysis of this data should include a correction for the 30ms stimulus delay mentioned previously in the description of the data collection methods for the facial emotion task.

Source Number	Color	Location
0	Blue	Left Inferior Parietal
1	Green	Left Occipital/Primary Visual
2	Red	Left Medial Superior Parietal
3	Cyan	Left Occipital/Primary Visual
4	Magenta	Left Inferior Temporal/Medial Wall/Amygdala
5	Yellow	Right Medial Superior Occipital
6	Black	Right Anterior Inferior Temporal/Medial Wall
7	Purple	Right Striate

Table 7: Facial Emotion Group Model Source Locations

Details of the facial emotion group model sources as seen in Figure 34. The primary (sources 1, 3) and subsequent visual areas (source 5) are well represented, as are sources which plausibly represent the left (source 0) and right striated cortex (source 7). Sources reaching forward into the left (source 4) and right inferior temporal regions (source 6) are also seen. Both inferior temporal sources are either approaching the medial side of the cortex (source 6, right), or fully on the medial wall (source 4, left). Source 4 in particular could be construed to represent the left amygdala, although this structure does not appear in a cortical model. Left parietal (source 2) and near parietal (source 0) sources not predicted by the Adolphs model are also seen.

Results of the group model suffer from the limitation that all twenty-three subjects were coregistered directly onto the average brain. This creates a level of imprecision in that the shapes of the cortical surfaces actually responsible for the original, measured activities do not match the assumed shape being used to model the subjects. This imprecision is assumed to create ambiguity with regard to source localizations, and the use of surrogate average brains is widely considered poor practice in MEG analysis. Individual MRIs are available for fifteen of the subjects, and single-subject analysis was performed on these fifteen using both their individual head models and the average. Goodness of fit for these models may be compared between columns 2 and 3 of Table 5. Comparison of these values reveals that use of the average brain may not be as detrimental as commonly assumed. The worst fit among these fifteen subjects, accounting for 8% of the variance of the original evoked response, occurs using an individual head model (subject ‘2’). The best fit accounts for 52% variance using both an individual head model and the average (subject ‘7’). Overall, the values seen in these columns are similar, with no obvious difference between individual head models and the average brain.

Values in column 4 of Table 5, representing the goodness of fit of the group model to the individual subjects, are seen to be uniformly smaller than either of columns 2 and 3. This indicates that the group model does not fit the individual subjects as well as their own, single-subject models. In only one case, that of subject ‘11’, does the group model even approach the individual model (51%, 49%, 43%). Some decrease in fit is the expected result of a group model, but the degree of the differences between individual and group model fits is disturbing. Several of the individual models account for nearly 50% of the original variance in their respective subjects, indicating that these subjects are modeled at least modestly well. With percentages of variance accounted for dipping as low as 1% for several subjects in the group model, the degree to which this model usefully represents these subjects must be questioned.

Fit of the group model is seen to be uneven between subjects as well, with the best fit approaching 20% of variance in several cases, but the worst being the aforementioned 1%. If the group model fit the individual subjects more uniformly, perhaps we might be reassured of its utility. Along these lines it is often observed in MEG analysis that the levels of both peak and overall measured power differs between subjects. Hypothesizing that the subjects poorly fit by the group model are those with lower overall power output, and therefore lesser influence on the group model which ultimately does not fit them, some type of between-subjects power normalization is indicated. The individual sources in the group model are broken out by fit within each subject in Table 6.

Limitations of this Work

The most immediate limitation of the software pipeline developed here is the inability to morph individual brain topology into a single, common source-space. Reliance on direct registration to an ill-fitting average head model is expected to severely weaken results. This is far from the only shortcoming of this pipeline but, without a technical resolution to this problem, credible work on Group-irMxNE cannot continue. Fortunately there are recent indications that the MNE developers are preparing to undertake a solution. Group-irMxNE is also appropriate for forming aggregate models of multiple response conditions within individual subjects (i.e. happy versus sad faces), and this functionality does not rely on morphing head models.

An important limitation inherent in the MxNE family of solvers in the context of this work is the way in which they deal with uneven power over time in the evoked response. Brain network models, such as that of Adolphs, are often information-centric rather than power-centric. Evoked responses, on the other hand, do not contain even power over time, leading to periods of high power being responsible for the placement of multiple sources at the expense of those with low power. It is likely that this is partially responsible for the difficulty placing frontal sources corresponding to later stages of the Adolphs model. The solution to this problem likely lies in some type of temporal variance normalization procedure, which may be applied to the original sensor-level trials data, rendering their power constant over time.

Evoked power levels are also observed to vary between subjects. The software pipeline developed here does not contain a provision to correct for this, and it is likely that those subjects with greater electromagnetic output will have disproportionate influence on Group-irMxNE models. A procedure is needed to normalize power between subjects prior to group modeling.

In this work, the selection of the regularization parameter (λ) input to irMxNE and Group-irMxNE is based on either prior knowledge of the actual number of sources responsible for the data under analysis (as in the simulations), or simply an arbitrary upper limit placed on the complexity of the system beyond which interpretation was deemed to become too difficult (eight sources in the facial emotion analysis). (As mentioned previously, the number of resamplings possible on the Granger spectra is constrained by computer memory limitations, and this constraint also becomes more severe as the number of sources increases.) A more principled method of selecting the number of sources included in the individual and group models is needed. Methods similar to those used to choose the number of components retained in principle components analysis suggest themselves. These methods are based upon the total variance accounted for by the model and the successive margins contributed by additional sources. (It is suspected but should be verified that the MxNE solvers generate source localizations ordered by decreasing variance.) On the other hand, as λ goes to zero the MxNE solution becomes identical to the L2-norm distributed solution. The phenomenon of splitting sources, similar to that seen in independent components analysis, should be guarded against.

The lack of a group inference procedure is, as discussed above, hampered by the unknown distribution of the measures of Granger influence. Without inference – the final step in most analyses – it is difficult to consider this pipeline complete. Additionally, reliance on resampling methods is cumbersome. If an alternative to resampling cannot be found then any improvements in the run-time and resource utilization of the Granger estimation program would certainly be appreciated by future users.

More general limitations of this work exist as well. It was discussed several times as to whether the neural response of multiple subjects would be sufficiently similar to make a group model possible. As opposed to fMRI-based group models, which are largely based on similar source localizations, a Group-irMxNE model also requires similar activity in time. It is the experience of the author that, although MEG responses are recognizably similar between subjects, they are not identical. Additionally, dissimilarity between subjects may be expected to become greater at longer post-stimulus times (as small differences in timing accumulate).

The localizations performed in this work are based upon the full bandwidth of the data (i.e. 0.2-55 Hz). This is desirable because it results in a single model, versus models for individual neural frequency bands (delta, theta, beta, gamma, etc.). However, such a wideband localization assumes that the power of all relevant frequencies is equally represented at the input to the localizer. This problem is shared with many neuroimaging procedures and is not unique to this work, but should remain in researchers' awareness as new methods are developed.

A pair of questionable assumptions are made by this analysis. First, MVAR analysis and the Fourier transform assume a linear model which, although generally counter-indicated for the brain, is supported for short-window analysis. Second, it is assumed that neural activity is constrained to the cortical surface. This is also a simplifying assumption which is not supported.

With regard to the facial emotion data collected here, the subjects represent a sample of convenience and are likely biased towards the demographics typical of a university campus, namely youth, health, higher socioeconomic status, intelligence, and educational achievement. The external validity of the limited conclusions drawn in this work based upon these subjects may be questioned on this basis. The facial evoked responses used here contain trials from seven facial affective conditions (happiness, sadness, anger, fear, surprise, disgust, and neutral). While this results in a large number of trials, and hopefully a high signal-to-noise ratio, only the portions of the neural response common to all faces are supported by this procedure. Greater specificity in terms of localization and connectivity may be possible by analyzing the facial conditions separately. By the same token, the subjects represent both genders and a wide range of ages, and gender-specific, age-specific analyses may be called for. On the other hand, a true exploratory analysis should be able to uncover these differences in a data-centric manner without the imposition of external constraints. Both points of view are worth keeping in mind.

Future Work

The significant, immediate challenges detailed in the previous section represent a large amount of future work on this analysis pipeline. Morphing, inter- and intra-subject power normalization, and group inference are all important. Simulations 5 and 6 indicate that results of the pipeline degrade as noise levels increase, and the effects of higher noise levels on localizations and connectivity spectra should be investigated. More theoretical questions, such as the benefits of narrowband localization, should be entertained as future results suggest. Most of the issues mentioned here have applicability beyond this pipeline, and addressing them is likely to advance the frontiers of MEG analysis as a whole. The question of the sufficiency of between-subjects commonality for group MEG analysis, remains unanswered. It may be necessary to investigate methods of creating group models which allow more flexibility than does Group-irMxNE.

During the course of this work, MEG data was collected on three subjects in the median nerve stimulation protocol. The intention was to begin to validate the pipeline on a neural model with known localizations and connectivity (Mauguiere, et al., 1997; Kiebel, David, & Friston, 2006; Sutherland & Tang, 2006), and which is simpler than the facial perception model of Adolphs (2002a; 2002b). Additionally, Gramfort, et al. have had good success localizing median nerve stimulation data (2011). 120 trials per subject were collected on an Elekta Neuromag 306 system, and the data was preprocessed using the same procedures used in the facial emotion analysis. A signal-space projection was used to minimize the stimulus artifact. Unfortunately, both individual and group localizations and Granger connectivity results were not as clean as in the facial emotion output, and therefore were not useful in the present work. It is likely that a combination of insufficient trials and ineffective means of eliminating the stimulus artifact are responsible for this. It would be beneficial to revisit the analysis of this protocol using publicly available median nerve stimulation data, with known localizations and connectivity estimates.

This work expends most of its energy developing the analysis pipeline, and makes only broad, general progress towards validating and refining the Adolphs model of facial emotion perception. Additional effort is needed in the analysis of this data, particularly once the question of source-space morphing is resolved. The right fusiform gyrus, containing the fusiform face area, shows considerable activation, and the placement of a source at this obvious location is of concern. Narrowband localization may be needed, as may time-windowing, and the specification of a more homogeneous group of stimulus conditions and subjects. However, it should be remembered that this pipeline represents an attempt at exploratory analysis, and that it attempts to avoid the need for such external constrains Overall, it is hoped that this pipeline lays the foundation for the validation of this and other neural connectivity models in the future at the group level. Independently of this analysis, the facial emotion data utilized here remains a remarkably clean and comprehensive MEG dataset which is worthy of continued attention.

Neuroimaging research is often driven by the availability of user-friendly software to implement new techniques. It is hoped that the pipeline developed here will encourage future group connectivity studies in MEG, and that awareness will be raised regarding the assumption of stationarity.

The success of this work in establishing connectivity using post-localization source timecourses points out that addressing many current neuroscience questions depends upon not only localizing neural sources, but also upon processing which takes place downstream from these localizations. The unprecedented success of irMxNE followed by the multivariate single-trials regression in producing high-fidelity source timecourses which enable such subsequent processing suggests that this be applied as a new standard to existing and future localization methods. As such it is indicated that current sparse localization and components analysis schemes should be evaluated for the degree to which they support follow-on analyses.

Summary and Conclusions

In recent years neuroimaging research has moved away from the simple identification of neural sources and has begun to address the dynamic systems formed by these sources. MEG is well positioned in this regard due to its excellent temporal resolution, but much MEG-based work continues to focus on source localization. When connectivity is estimated, Granger causality is frequently employed, but almost exclusively in a form based on MVAR estimation which depends upon stationarity and is unsuitable for dynamic systems. Additionally, connectivity analyses are almost always conducted at the level of individual subjects and have difficulty generalizing beyond the idiographic level. One solution lies in developing appropriate techniques for group connectivity analysis.

The ambitious goal undertaken in this work is that of establishing a software pipeline for the group, non-parametric Granger causality analysis of MEG subjects. The irMxNE sparse localizer found in the MNE-Python software belongs to a family of solvers specifically developed to work with non-stationary data. A group extension of this localizer is created, and a multivariate multiple regression procedure is established to generate single-trial output. These single trials are submitted to a custom implementation of the complex Morlet wavelet designed to maintain constant peak wavelet power across frequencies. The resulting wavelet spectra are passed to a wrapper program constructed in Matlab around the non-parametric, conditional Granger causality functionality accessed through the *ft_connectivity_analysis* interface of the Fieldtrip package. This wrapper provides the missing support for the time-frequency domain as well as enables parallel computation of results. Resampling-based inference is provided on individual Granger connectivity spectra output from this pipeline, but a procedure to provide group inference has not been identified. Other outstanding technical issues remain as well, including morphing of multiple subjects’ head topology into a common source-space, normalizing power between and within subjects, and choosing a value of the regularization parameter.

A pair of key issues in neuroimaging analysis are highlighted by this work. First, by utilizing irMxNE and non-parametric Granger causality, this pipeline avoids the common and unwarranted assumption of stationarity. This assumption precludes the analysis of dynamic systems, and therefore it is important to identify techniques not bound by stationarity. Second, analyses often focus on source localization, with only secondary consideration given to timecourses. In performing connectivity analysis on post-localization source timecourses, this work illustrates the importance of localization techniques which faithfully deliver timecourse data suitable for follow-on analysis.

Single-subject simulations based on a well-known, bivariate, non-stationary MVAR system result in the original source localizations being returned and the Granger influence measures of the original system being replicated. This validates the end-to-end functionality of the entire pipeline for single subjects, including source localization, single-trials propagation, and Granger causality estimation.

Multi-subject simulations suggest that the pipeline is well able to separate patterns of sources and connectivity which are clearly distinct between subjects. Results also suggest that, when multiple subjects contain similar patterns, the pipeline attempts to generalize them into a true, group solution.

Source localizations of both individual facial emotion subjects and the subjects as a group are hindered by irMxNE’s handing of uneven power over the duration of the evoked response. Visual processing, radiating forward through the occipital, parietal, and temporal regions during the period 100-200 ms, and which generates a large amount of power, is well represented with numerous source localizations. These areas correspond to the early processing stages of the Adolphs model. Later stages, occurring during times of less overall power, including processing in the inferior, orbito-frontal, and somatosensory cortices, are not localized.

Connectivity estimates between sources in both individual and group output show a pattern which simulation suggests should be expected in a realistic neural system. This pattern consists of relatively low frequency connectivity at varying latencies following stimulus presentation. Connectivity spectra developed here includes the pre-stimulus interval and, although limited and seemingly random results are seen here as well, overall results indicate that both individual and group analyses are identifying actual connectivity in cortical systems.

Localizations and connectivity estimates for individual subjects are idiographic, and no attempt is made to interpret them. For the group model, localizations generally support early stages of the Adolphs model.

In conclusion, it is believed that this work represents an important contribution to electromagnetic neuroimaging analysis. The group, non-stationary MEG connectivity analysis pipeline suffers from a handful of outstanding technical issues. These issues prevent its performance on actual groups of MEG subjects from being fully assessed. Nonetheless, and regardless of the ultimate success of this particular combination of techniques, it is hoped that this pipeline represents the type of effort needed to address a new generation of post-localization neuroscience questions.

References

Adolphs, R. (2002a). Recognizing emotion from facial expressions: Psychological and neurological mechanisms. *Behavorial and Cognitive Neurosceince Reviews, 1*, 21-61.

Adophs, R. (2002b). Neural systems for recognizing emotion. *Current Opinion in Neurobiology, 12*, 169-77.

Baillet, S., Mosher, J. C., & Leahy, R. M. (2001). Electomagnetic Brain Mapping. *IEEE Signal Processing Magazine, 18*(6), 14-30.

Barbour, A. J., & Parker, R. L. (2014, April 15). *Normalization of Power Spectral Density Estimates*. Retrieved from cran.r-project.org: cran.r-project.org/web/packages/psd/vignettes/normalization.pdf

Barnett, L., & Seth, A. K. (2011). Behaviour of Granger causality under filtering: Theoretical invariance and practical application. *Journal of Neuroscience Methods, 201*, 404-19.

Barnett, L., & Seth, A. K. (2014b). The MVGC multivariate Granger causality toolbox: a new approach to Granger-causal inference. *Journal of Neuroscience Methods, 223*, 50-68.

Bassett, D. S., & Lynall, M. E. (2013). Network Methods to Characterize Brain Structure and Function. In *Cognitive Neurosciences: The Biology of the Mind* (pp. 1-27).

Bell, A. J., & Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural Computatoin, 7*, 1129-59.

Calder, A. J., & Young, A. W. (2005). Understanding the recognition of facial identity and facial expression. *Nature Reviews Neuroscience, 6*, 641-51.

Calhoun, V. D., Adali, T., Pearlson, G. D., & Pekar, J. J. (2001). A Method for Making Group Inferences from Function MRI Data Using Independent Component Analysis. *Human Brain Mapping, 14*, 140-51.

Calhoun, V. D., Adali, T., Pearlson, G. D., & Pekar, J. J. (2001). A method for making group inferences from functional MRI data using independent component analysis. *Human Brain Mapping*, 140-51.

Calhoun, V. D., Adali, T., Pearlson, G. D., & Pekar, J. J. (2001). Spatial and Temporal Independent Component Analysis of Functional MRI Data Containing a Pair of Task-Related Waveforms. *Human Brain Mapping, 13*, 43-53.

Cardoso, J. F. (1999). Higher-order contrasts for independent component analysis. *Neural Computation, 11*, 157-92.

Collura, T. F. (2008). Towards a coherence view of brain connectivity. *Journal of Neurotherapy, 12*, 99-110.

Correa, N., Adali, T., & Calhoun, V. D. (2007). Performance of Blind Source Separation Algorithms for FMRI Analysis using a Group ICA Method. *Magnetic Resonance Imaging, 25*(5), 684-94.

Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage, 9*, 179-194.

Darwin, C. (1965). *The Expression of the Emotions in Man and Animals* (5th ed.). Chicago: University of Chicago Press.

David, O., Guillemain, I., Sailet, S., Reyt, S., Deransart, C., Segebarth, C., & Depaulis, A. (2008). Identifying neural drivers with functional MRI: an electrophysiological validation. *PLoS Biology, 6*, 2683-97.

Delorme, A., Palmer, J., Oostenveld, R., Onton, J., & Makeig, S. (2007). *Comparing Results of Algorithms Implementing Blind Source Separation of EEG Data.* Swartz Foundation and NIH Grant.

Dhamala, M., Rangarajan, G., & Ding, M. (2008a). Estimating Granger causality from Fourier and wavelet transforms of time series data. *Physical Review Letters, 100*, 018701-1-4.

Dhamala, M., Rangarajan, G., & Ding, M. (2008b). Analyzing information flow in brain networks with nonparametric Granger causality. *NeuroImage, 41*, 354-62.

Dien, J. (1998). Addressing misallocation of variance in principal components analysis of event-related potentials. *Brain Topography, 11*, 43-55.

Dien, J. (2010). Evaluating two-step PCA of ERP data with Geomin, Infomax, Oblimin, Promax, and Varimax rotations. *Psychophysiology, 47*, 170-83.

Ding, M., Bressler, S. L., Yang, W., & Liang, H. (2000). Short-window spectral analysis of cortical event-related. *Biological Cybernetics, 83*, 35-45.

Ding, M., Chen, Y., & Bressler, S. L. (2006). Granger causality: basic theory and application to neuroscience. In *Handbook of Time Series Analysis: Recent Theoretical Developments and Applications* (p. 437).

Ding, M., Chen, Y., & Bressler, S. L. (n.d.). Granger causality: basic theory and application to neuroscience. In *Handbook of Time Series Analysis: Recent Theoretical Developments and Applications* (p. 437).

Eichele, T., Rachakonda, S., Brakedal, B., Eikeland, R., & Calhoun, V. D. (2011). EEGIFT: Group Independent Component Analysis for Event-Related EEG Data. *Computational Intelligence and Neuroscience*, 1-9.

Ekman, P. (1972). Universals and cultural differences in facial expressions of emotion. *Nebraska Symposium on Motivation* (pp. 207-283). Lincoln, NE: University of Nebraska Press.

Ekman, P. F., O'Sullivan, M., Chan, A., Diacoyanni-Tarlatzis, I., Heider, K., Krause, R. A., . . . Tzavaras, A. (1987). Universals and cultural differences in the judgements of facial expressions of emotion. *Journal of Personality and Social Psychology, 53*, 712-17.

Ekman, P., & Freisen, W. V. (1986). A New Pan-Cultural Facial Expression of Emotion. *Motivation and Emotion, 10*(2), 159-168.

Ekman, P., & Friesen, W. V. (1976). *Pictures of Facial Affect.* Palo Alto: Consulting Psychologists Press.

Friston, K. (2011). Functional and effective connectivity: a review. *Brain Connectivity, 1*, 13-36.

Friston, K., Moran, R., & Seth, A. K. (2013). Analysing connectivity with Granger causality and dynamic causal modeling. *Current Opinions in Neurobiology, 23*, 172-8.

Geweke, J. (1982). Measurement of linear dependence and feedback between multiple time series. *Journal of the American Statistical Association, 77*, 304-13.

Goupillaud, P., Grossmann, A., & Morlet, J. (1984). Cycle-octave and related transforms in seismic signal analysis. *Geoexploration, 23*(1), 85-102.

Gramfort, A., Kowalski, M., & Hamalainen, M. (2012). Mixed-norm estimates for the M/EEG inverse problem using accelerated gradient methods. *Physics in Medicine and Biology, 57*(7), 1937-61.

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., . . . Hamalainen, M. (2013). MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience, 7*, 1-13.

Gramfort, A., Luessi, M., Larson, E., Engermann, D., Strohmeier, D., Brodbeck, D., . . . Hamalainen, M. (2014). MNE software for processing MEG and EEG data. *NeuroImage, 86*, 446-60.

Gramfort, A., Strohmeier, D., Haueisen, J., Hamalainen, M. S., & Kowalski, M. (2013). Time-frequency mixed-norm estimates: sparse M/EEG imaging with non-stationary source activations. *NeuroImage, 70*, 410-22.

Gramfort, A., Strohmeier, D., Haueisen, J., Hamaleinen, M., & Kowalski, M. (2011). Funtional brain imaging with M/EEG using structured sparsity in time-frequency dictionaries. *Lecture notes in computer science, 6801*, 600-611.

Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica, 37*, 424-38.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences, 4*, 223-33.

Haykin, S. (1989). *An Introduction to Analog & Digital Communications.* New York: John Wiley & Sons.

Himberg, J., & Hyvarinen, A. (2003). Icasso: Software for investigating the reliability of ICA estimates by clustering and visualization. *Procedings of the 2003 IEEE workshop on neural networks for signal processing.*

Hutchison, R. M., Womelsdorf, T., Allen, E. A., Calhoun, V. D., Corbetta, M., Della Penna, S., . . . Chang, G. (2013). Dynamic functional connectivity: promise, issues, and interpretations. *NeuroImage, 80*, 360-78.

Kaplan, A. Y., Fingelkurts, A. A., Fingelkurts, A. A., Borisov, S. V., & Darkhovsky, B. S. (2005). Nonstationary nature of brain activity as revealed by EEG/MEG: Methodolgy, practical and conceptual challenges. *Signal Processing, 85*, 2190-212.

Kiebel, S. J., David, O., & Friston, K. J. (2006). Dynamic causal modeling of evoked responses in EEG/MEG with lead field parameterization. *NeuroImage*, *30*, 1273-84.

Klonowski, W. (2009). Everything you wanted to ask about EEG but were afraid to get the right answer. *Nonlinear Biomedical Physics, 3*, 1753-4631-3-2.

Lee, T.-W. (2000). A Unifying Information-Theoretic Framework for Independent Component Analysis. *Computers and Mathematics with Applications*, *39*, 1-21.

Liu, Y., Liang, X. S., & Weisberg, R. H. (2007). Rectification fo the bias in the wavelet power spectrum. *Journal of Atmospheric and Oceanic Technology, 24*, 2093-102.

Lysne, P. A. (2009). *An MEG Investigation of the Differential Responsivity of the Human Calcarine Fissure and Fusiform Gyrus to the Emotion of Viewed Faces.* Albuquerque, NM: University of New Mexico.

Lysne, P. A. (2014). *Controlling for Multiple Testing in Neuroimaging Analysis.* Comprehensive Examination Question, University of New Mexico, Psychology, Quantitative Area, Albuquerque.

Makeig, S., Bell, A. J., Jung, T.-P., & Sejnowski, T. J. (1996). Independent Component Analysis of Electoencephlographic Data. In D. Touretzky, M. Mozer, M. Hasselmo, & (Eds.), *Advances in Neural Information Processing Systems 8* (pp. 145-51). Cambridge: MIT Press.

Makeig, S., Jung, T.-P., Bell, A. J., Ghahremani, D., & Sejnowski, T. (1997). Blind separation of auditory event-related brain responses into independent components. *Neurobiology*, *94*, 10979-84.

Mauguiere, F., Merlet, I., Forss, N., Vanni, S., Jousmaki, V. A., & Hari, R. (1997). Activation of a distributed somatosensory cortical network in the human brain. A dipole modeling study of magnetic fields evoked by median nerve stimulation. Part 1: location and activation timing of SEF sources. *Electroencephlography and Clinical Neurophysiology/Evoked Potentials Section, 104*(4), 281-9.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Computational Intelligence and Neuroscience*.

Papoulis, A. (1991). *Probability, Random Processes, and Stochastic Processes, Third Edition.* New York: McGraw-Hill.

Park, H., & Friston, K. (2013). Structural and functional brain networks: from connections to cognition. *Science, 342*, 12384-11-1-8.

Richardson, M. H. (1978, March). Fundamentals of the Discrete Fourier Transform. *Sound & Vibration Magazine*, 1-8.

Rokem, A. T., & Perez, M. (2009). Nitime: time-series analysis for neuroimaging data. *Proceedings of the 8th Python in Science conference*, (pp. 68-75). Pasadena, CA.

Rubinov, M., & Sporns, O. (2010). Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage, 52*, 1059-69.

Schlogl, A. (2006). A comparison of multivariate autoregressive estimators. *Signal Processing, 86*, 2426-9.

Schumway, R. H., & Stoffer, D. A. (2006). *Time Series Analysis and Its Applications With R Examples, Third Edition.* New York: Springer Science+Business Media.

Seabold, S., & Perktold, J. (2010). Statsmodels: Econometric and Statistical Modeling with Python. *Procedings of the 9th Python in Science Conference*, (pp. 57-61). Austin.

Seth, A. K. (2010). A MATLAB toolbox for Granger causal connectivity analysis. *Jourmnal of Neuroscience Methods, 186*, 262-73.

Shumway, R. H., & Stoffer, D. S. (2006). *Time Series Analysis and Its Applications With R Examples Third Edition.* New York: Springer Science+Business Media.

Slepian, D., & Pollak, H. O. (1961, January). Prolate spheroidal wavefunctions Fourier analysis and uncertainty - I. *Bell Systems Technical Journal*, 43-63.

Stevens, J. P. (2009). *Applied Multivariate Statistics for the Social Sciences, Fifth Edition.* New York: Routledge, Taylor & Francis Group.

Strohmeier, D. (2014). Improved MEG/EEG source localization with reweighted mixed-norms. *Pattern Recognition in Neuroimaging, 2014 International Workshop On*, (pp. 1-4). Tubidgen.

Strohmeier, D., Haueisen, J., & Gramfort, A. (2014). The iterative reweighted Mixed-Norm Estimate for spatio-temporal MEG/EEG source reconstruction. *Preprint*.

Sutherland, M. T., & Tang, A. C. (2006). Reliable detection of bilateral activation in human primary somatosensory cortex by unilateral median nerve stimulation. *NeuroImage, 33*, 1042-54.

Thompson, D. J. (1982). Spectrum estimation and harmonic analysis. *Proceedings of the IEEE, 70*(9), 1055-96.

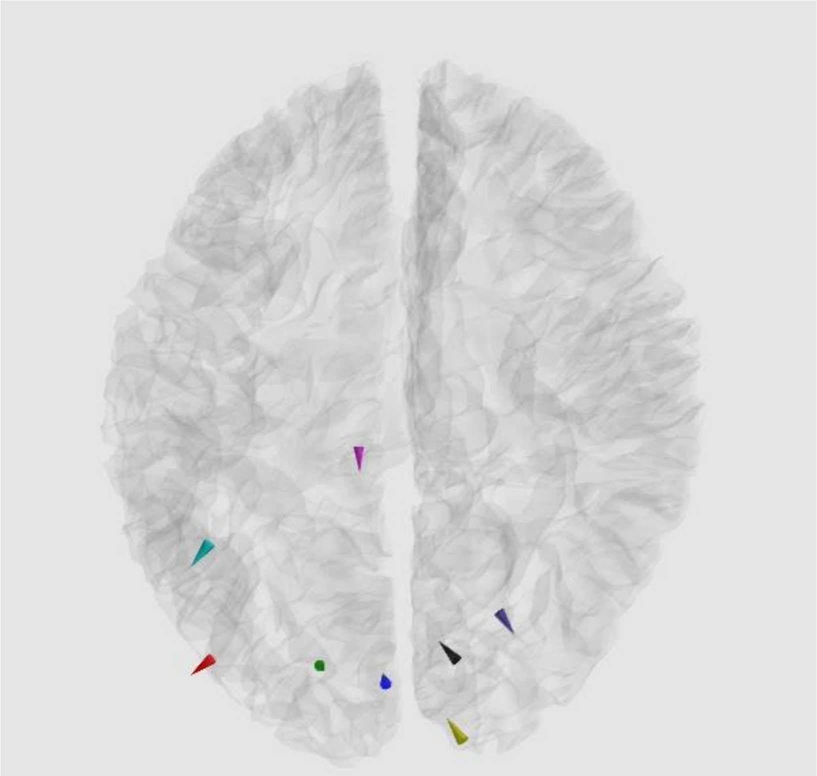
Torrence, C., & Compo, G. P. (1998). A practical guide to wavelet analysis. *Bulletin of the American Meterological Society*, *79*, 61-78.

Veleda, D., Montagne, R., & Araujo, M. (2012). Cross-wavelet bias corrected by normalizing scales. *Journal of Atmospheric and Oceanic Technology*, *29*, 1401-8.

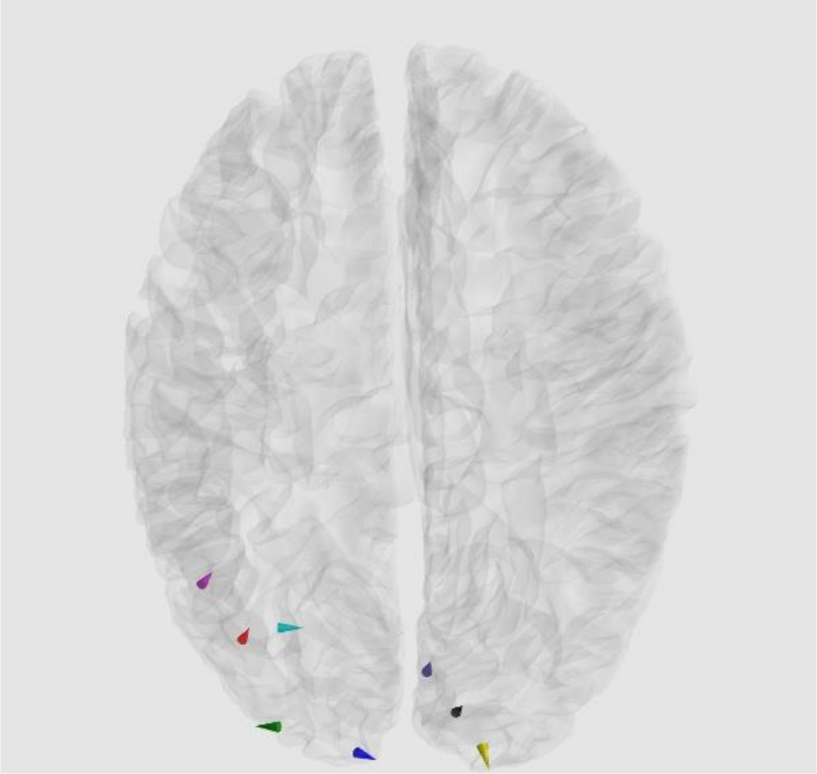
Wilson, G. T. (1972). The factorization of matrical spectral densities. *SIAM*, *25*, 420-6.

Wilson, G. T. (1978). A convergence theorem for spectral factorization. *Journal of Multivariate Analysis*, *8*(2), 222-32.

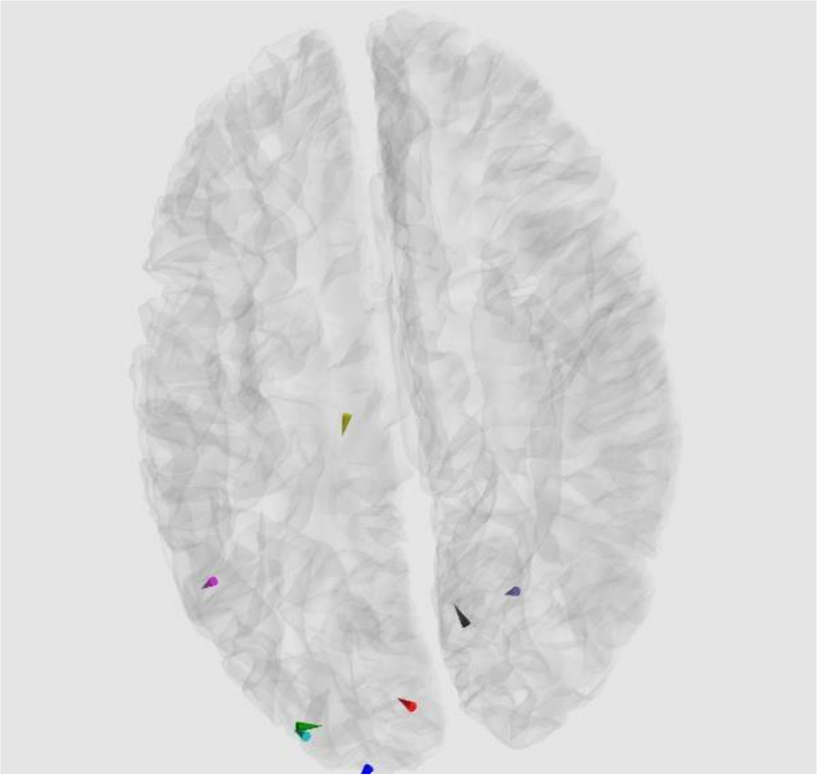
Appendix A: Facial Emotion Individual Subjects' Source Localizations



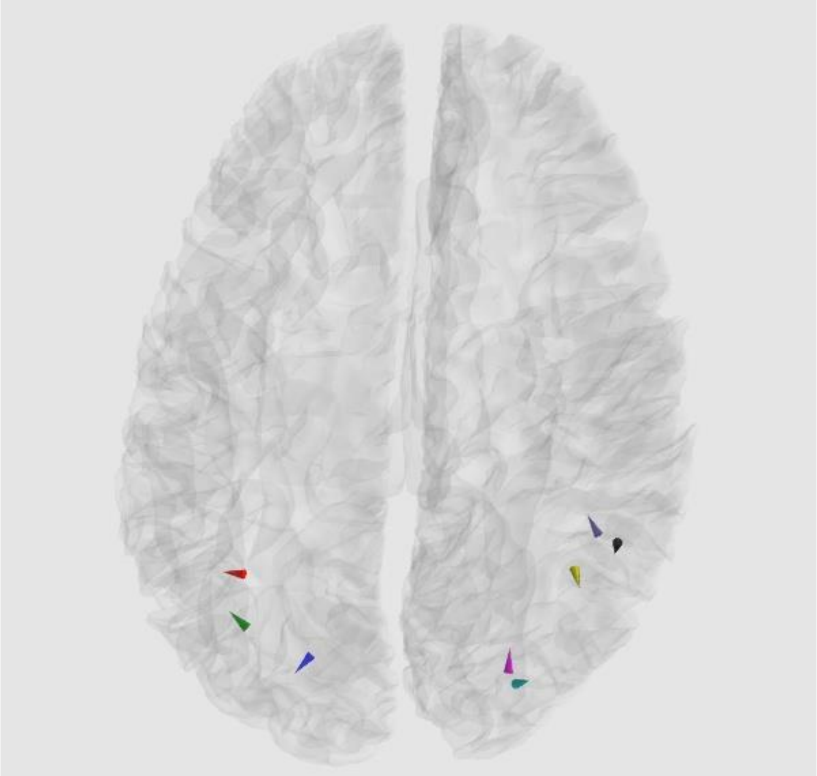
Subject 0



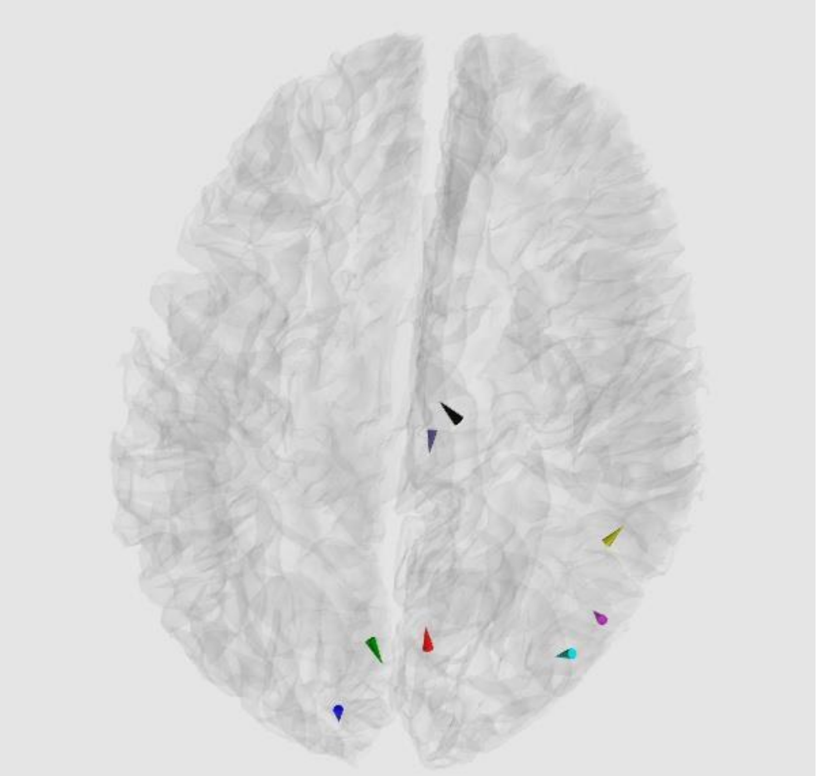
Subject 1



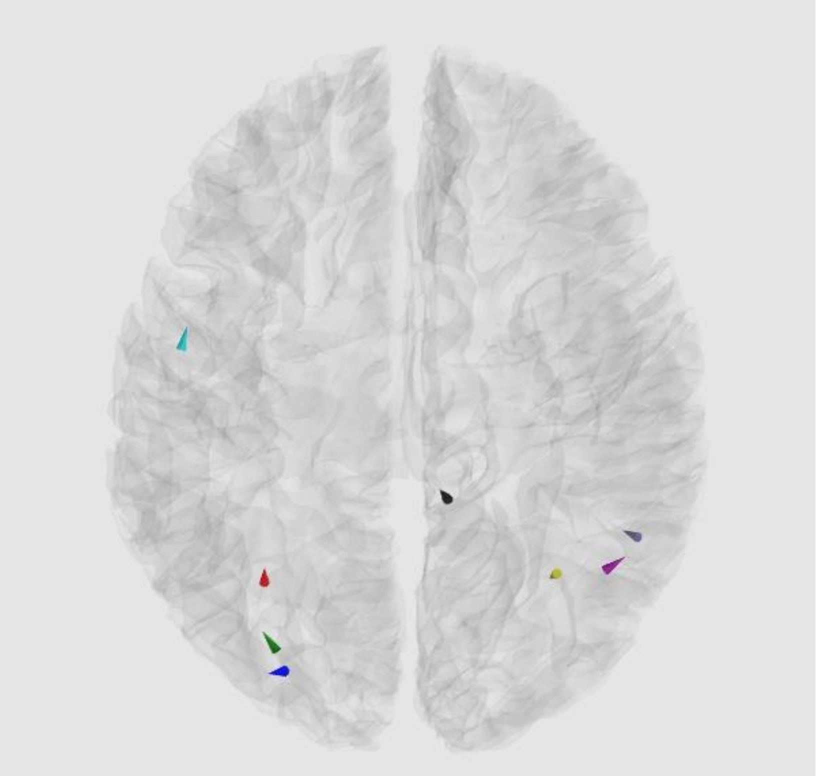
Subject 2



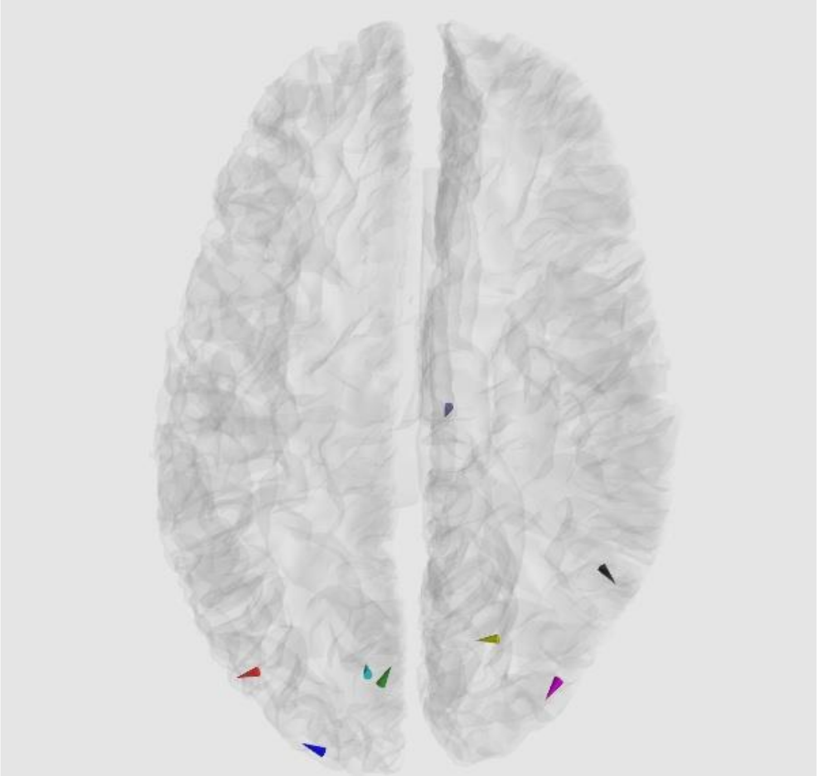
Subject 3



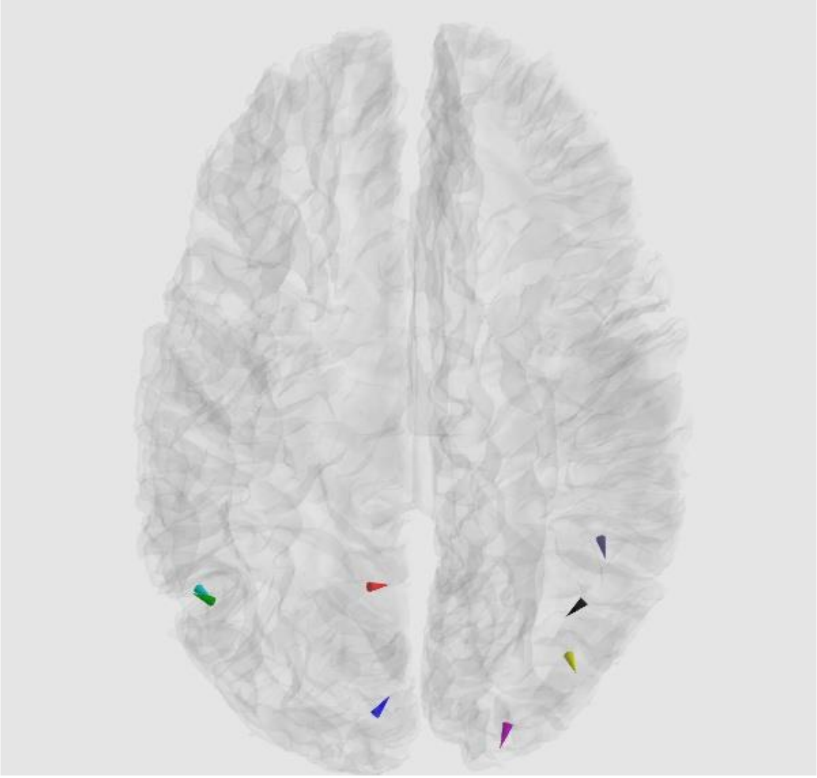
Subject 4



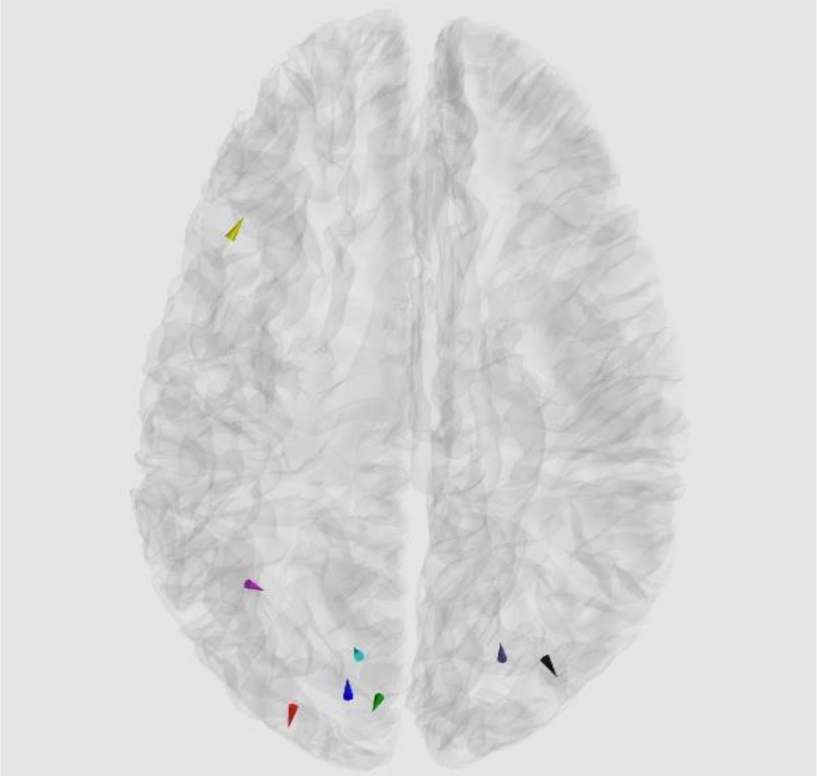
Subject 5



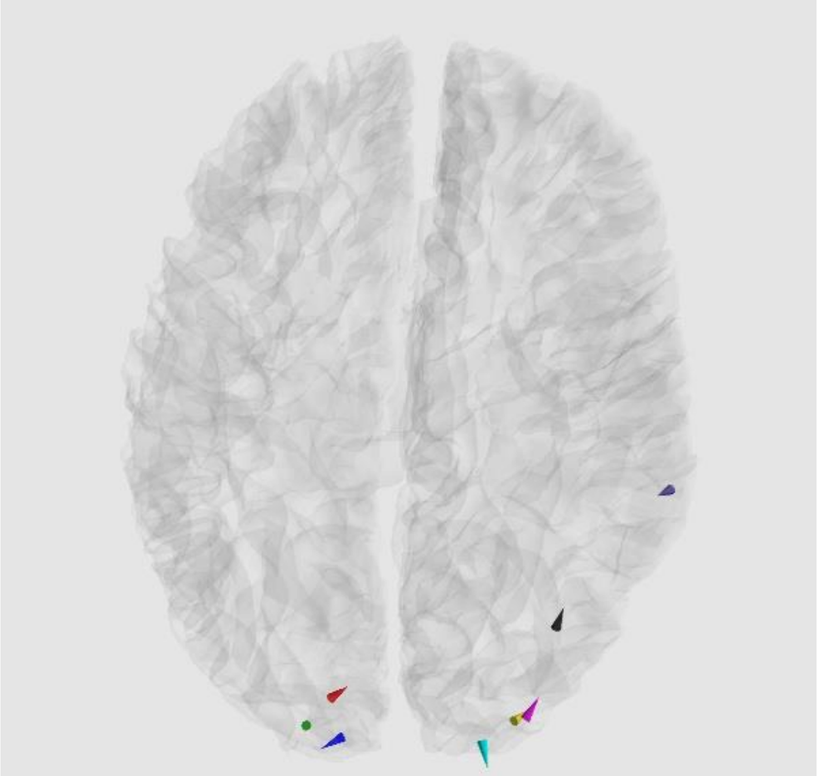
Subject 6



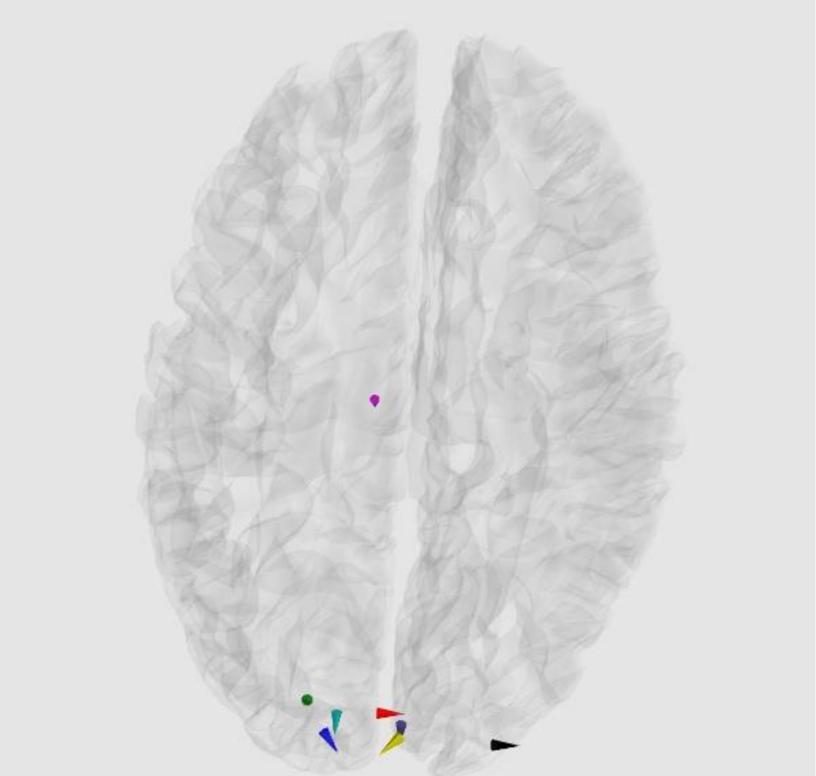
Subject 7



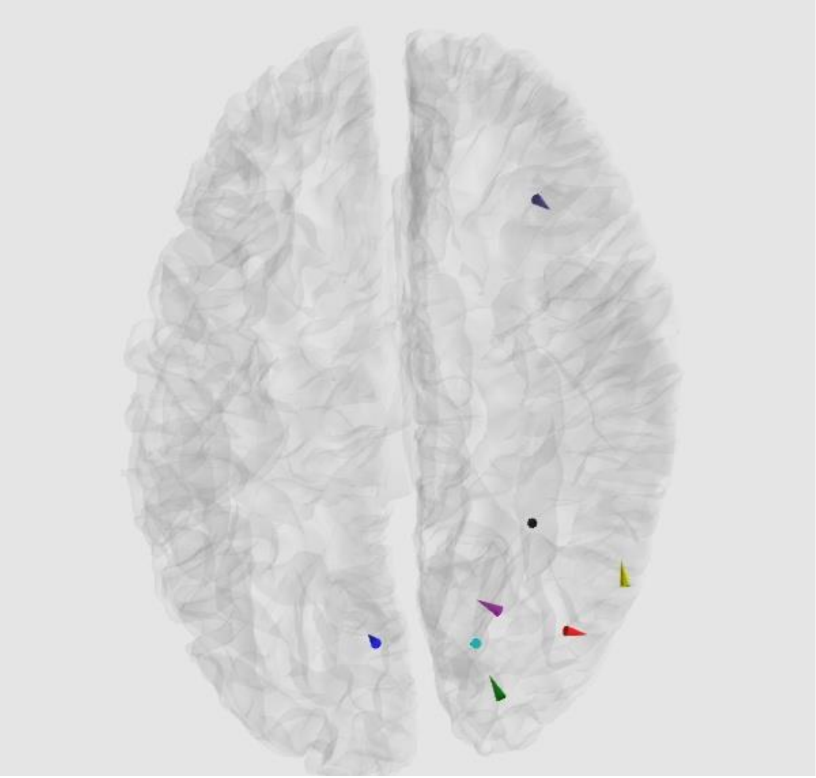
Subject 8



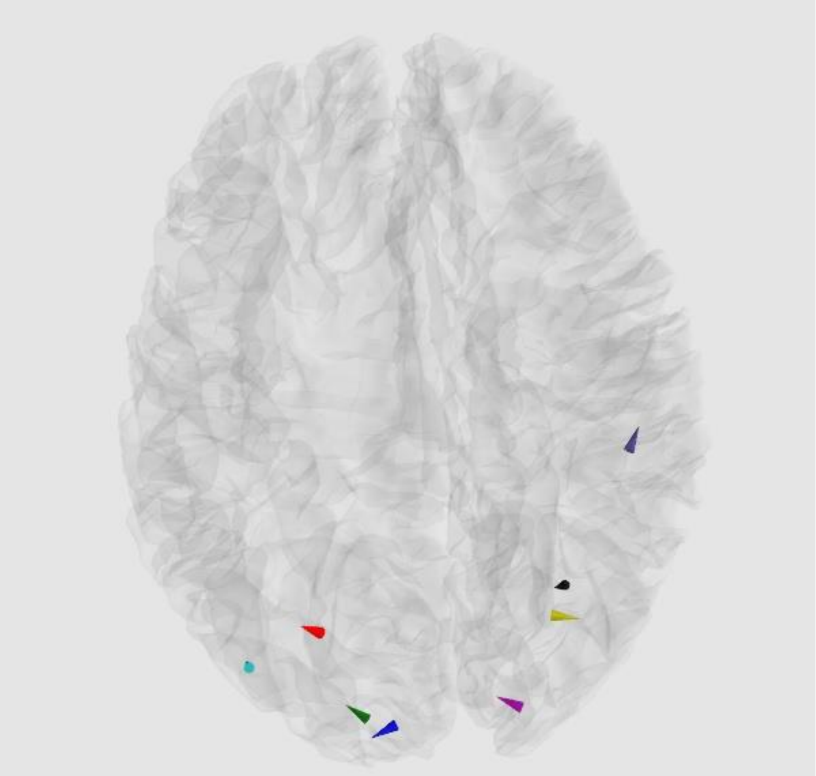
Subject 9



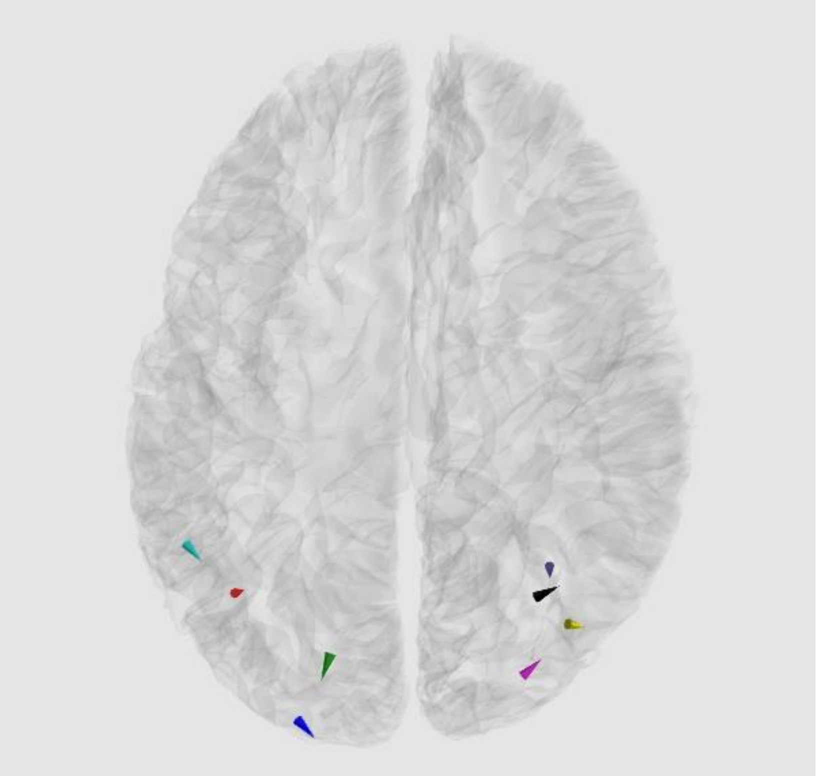
Subject 10



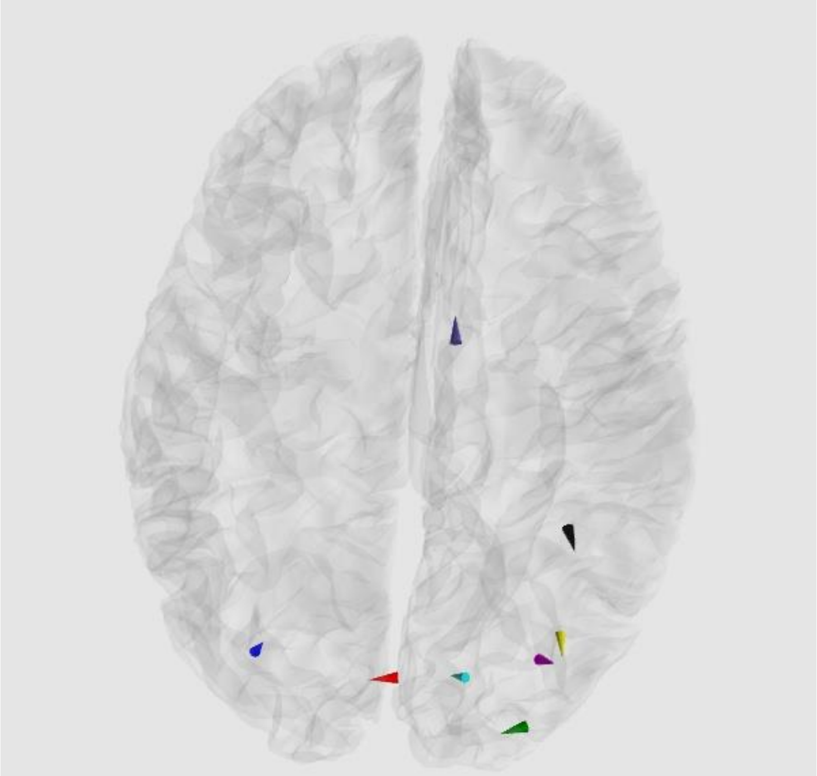
Subject 11



Subject 12



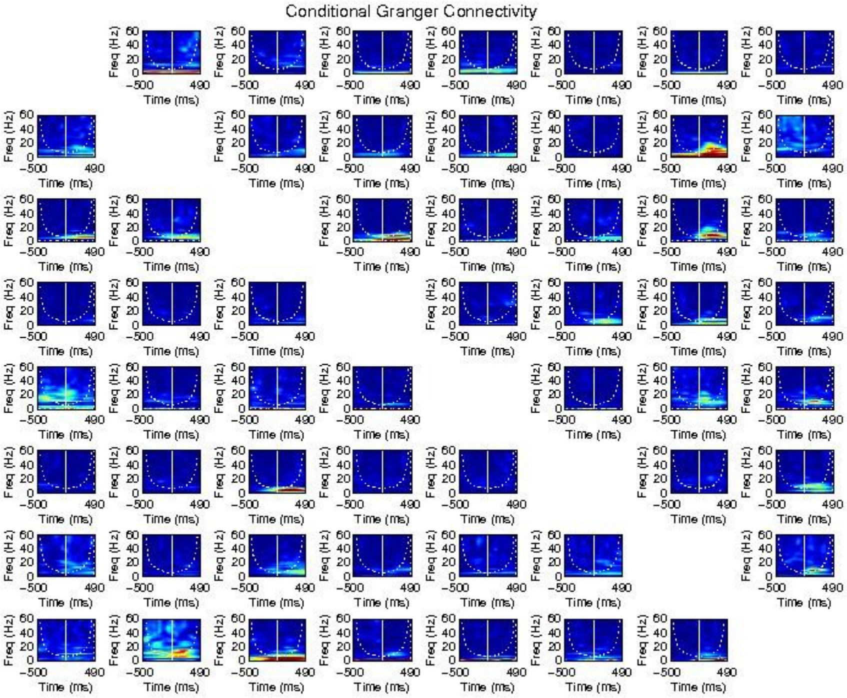
Subject 13



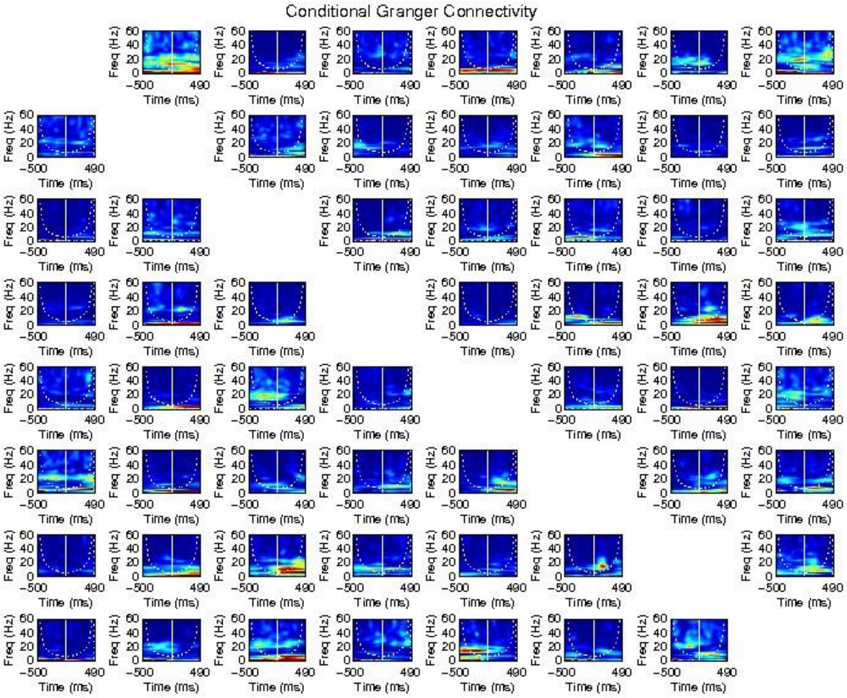
Subject 14

Figure A1: Facial Emotion Individual Subjects' Localizations

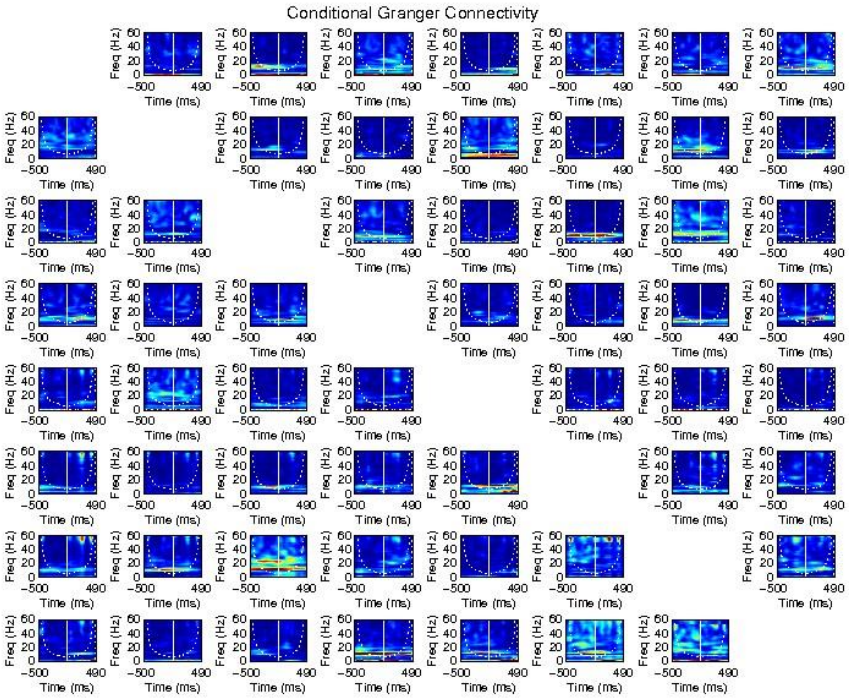
Appendix B: Facial Emotion Individual Subjects' Granger Spectra



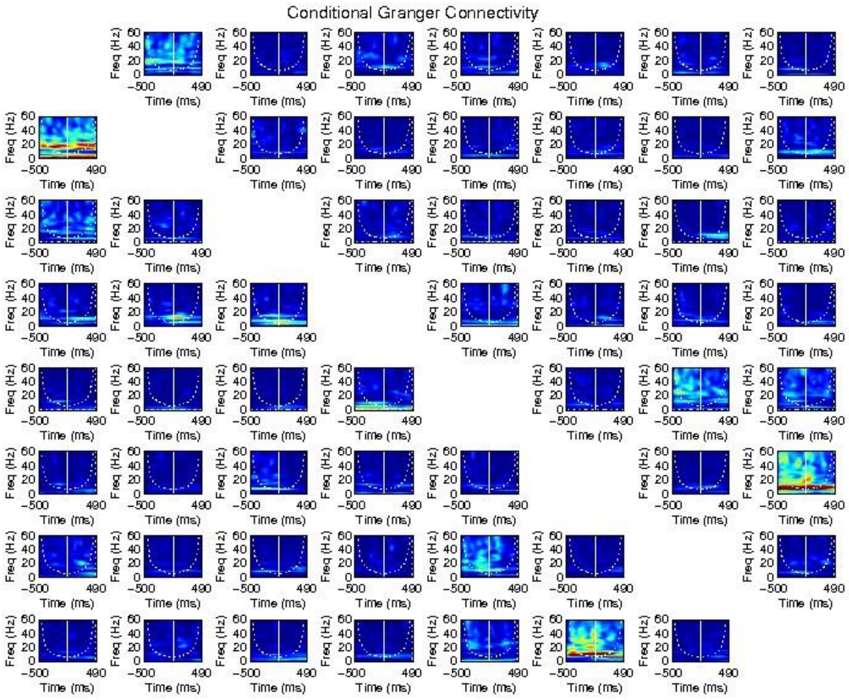
Subject 0



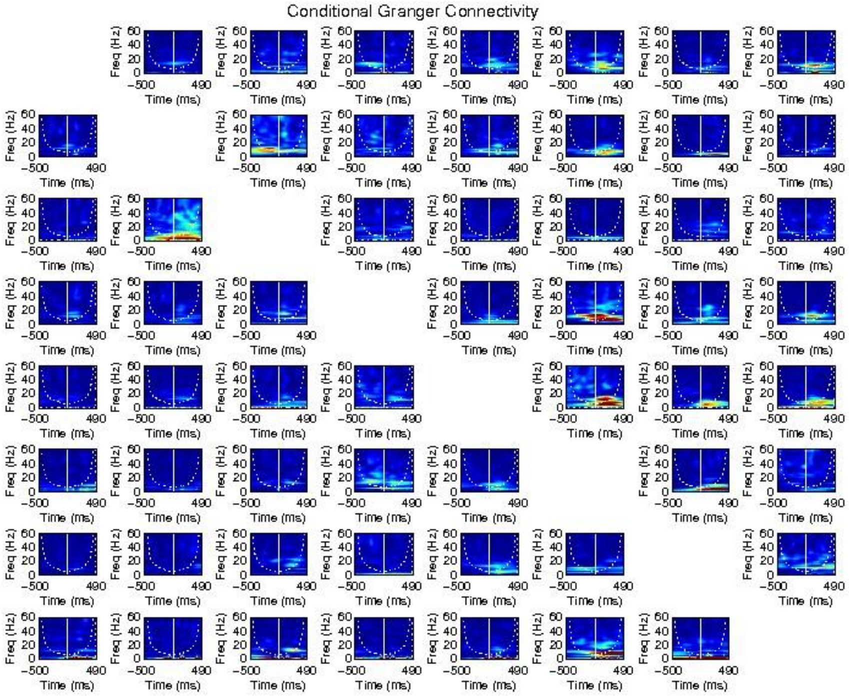
Subject 1



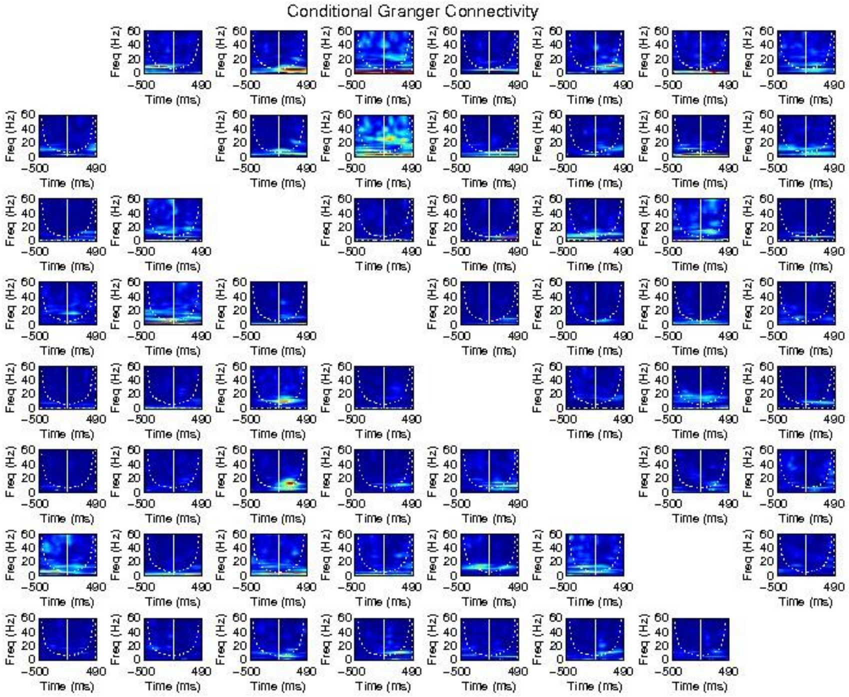
Subject 2



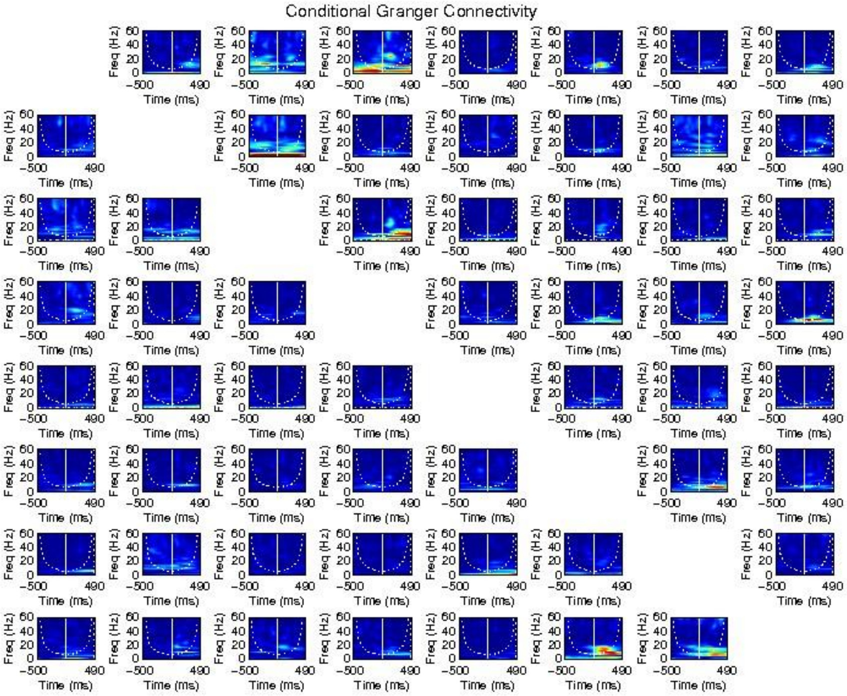
Subject 3



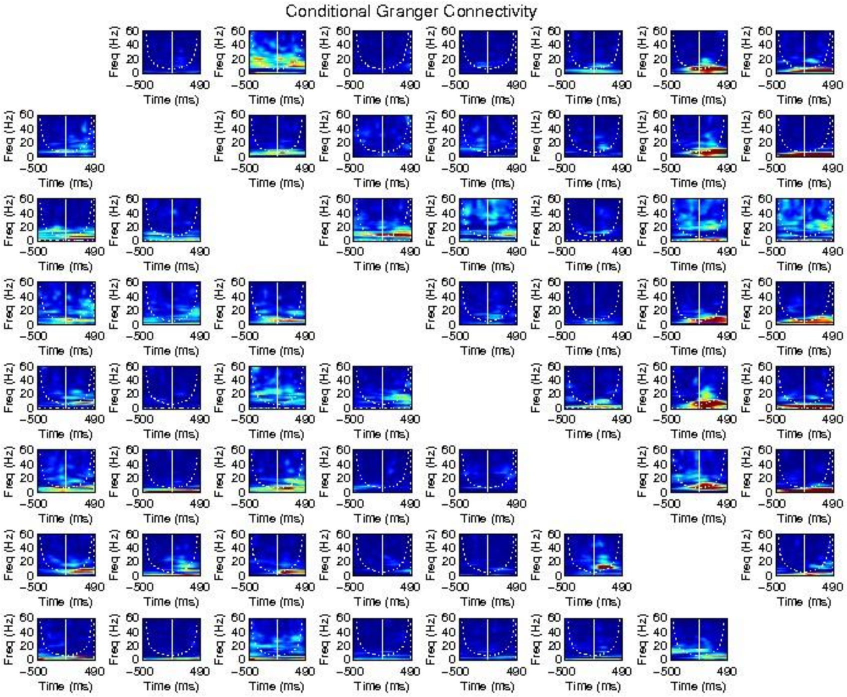
Subject 4



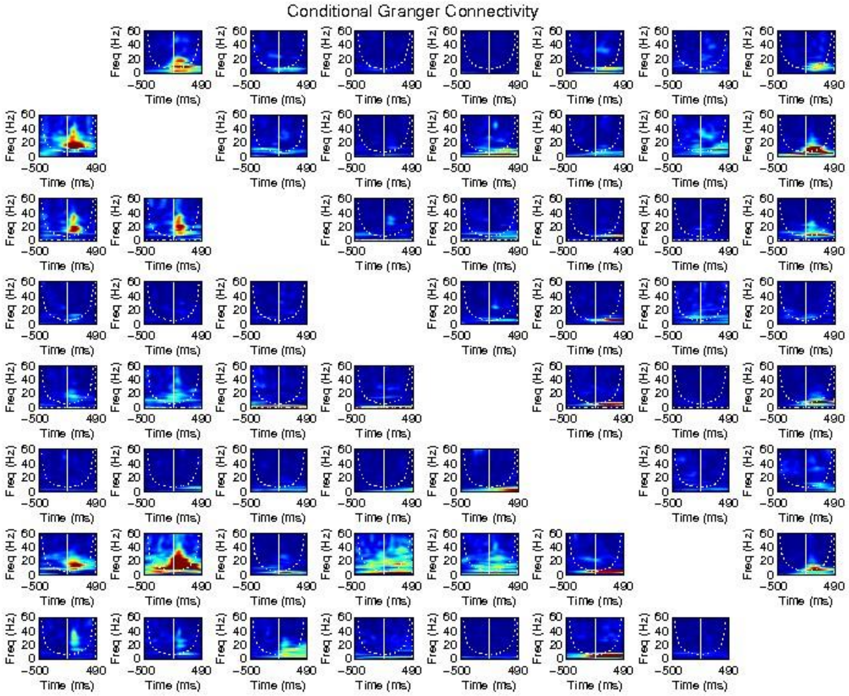
Subject 5



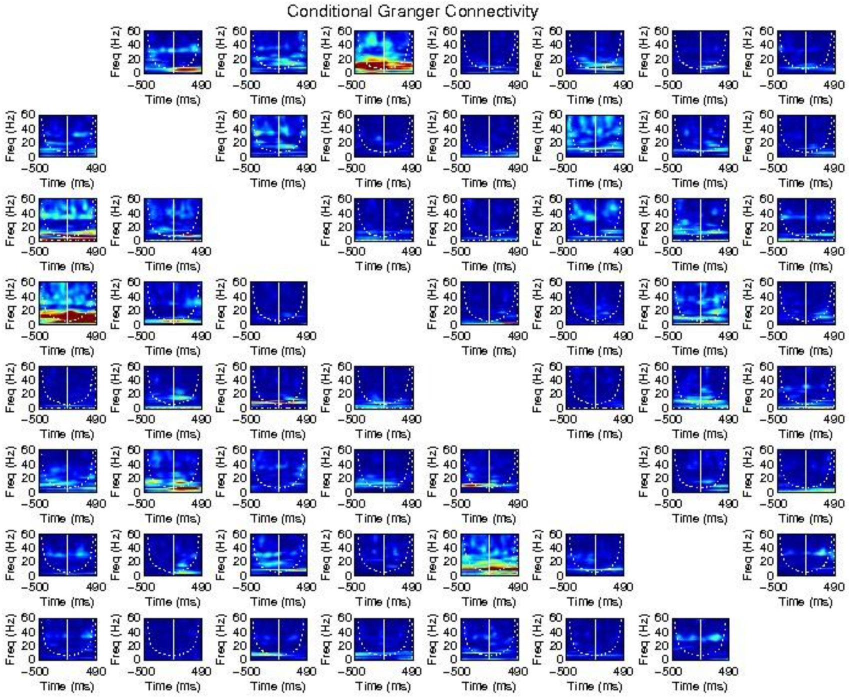
Subject 6



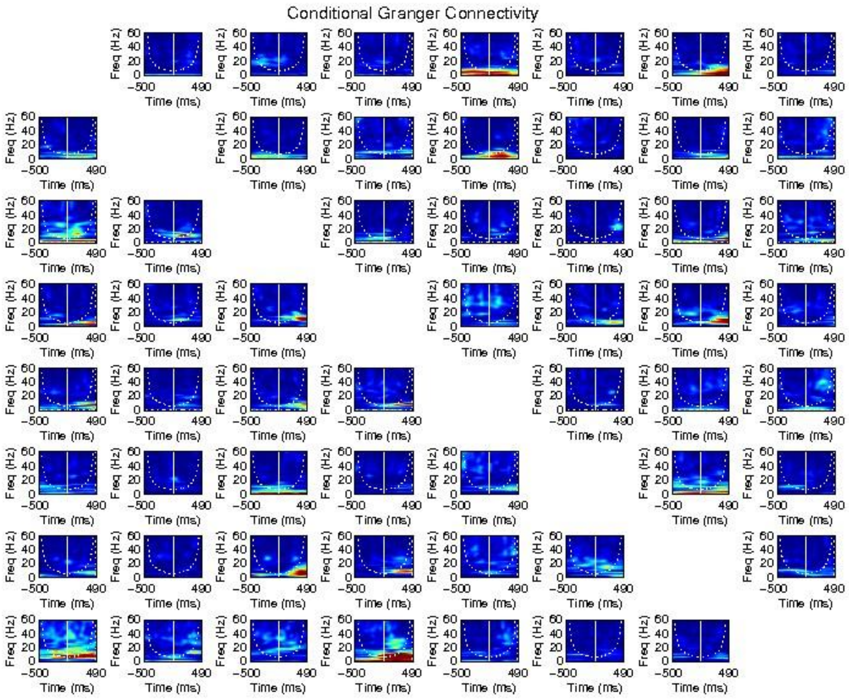
Subject 7



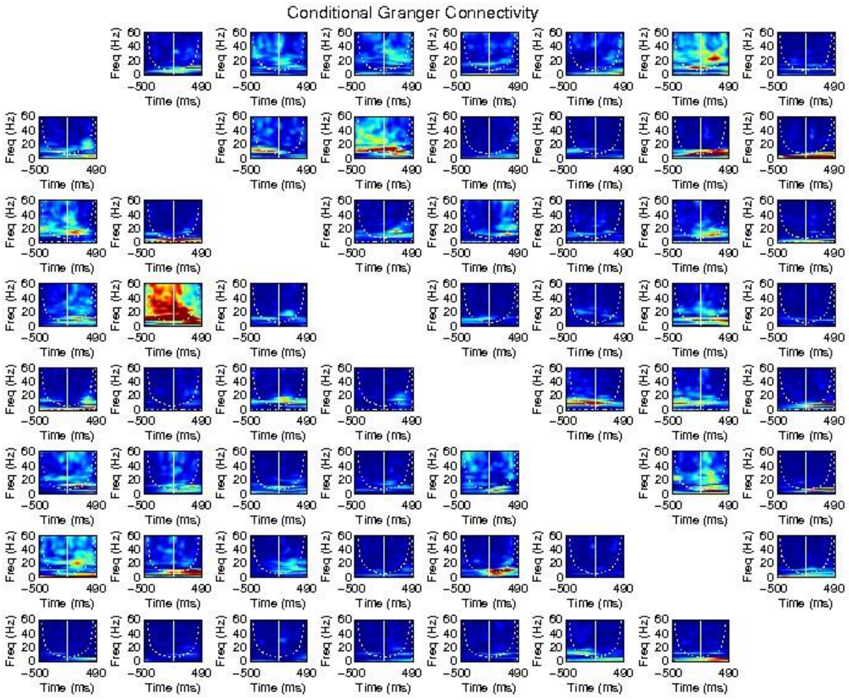
Subject 8



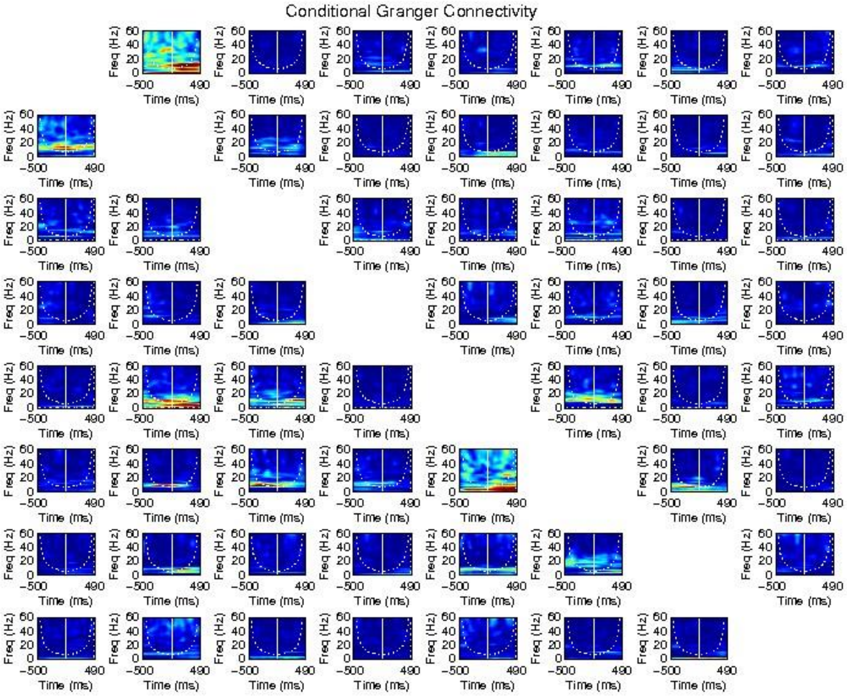
Subject 9



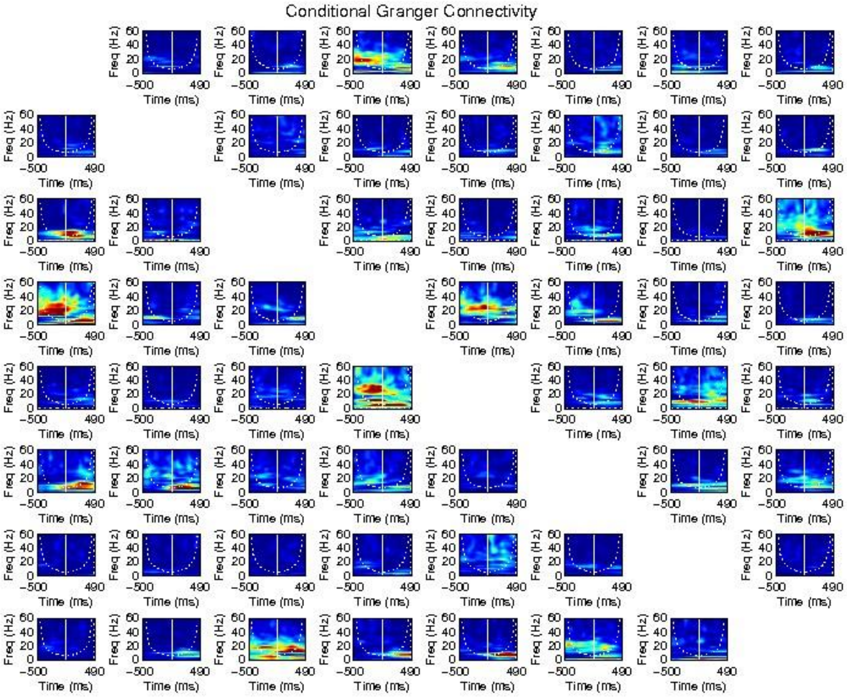
Subject 10



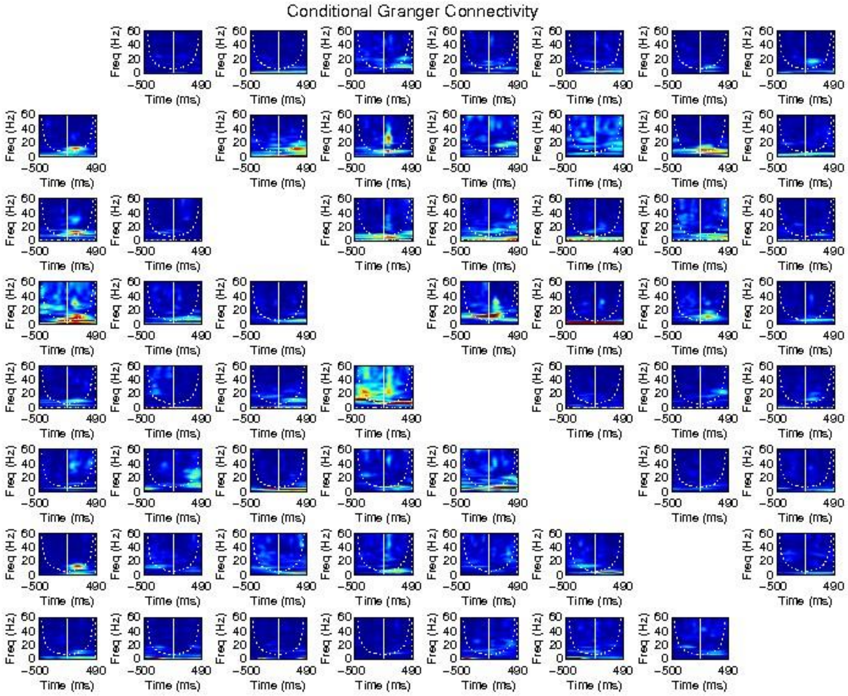
Subject 11



Subject 12



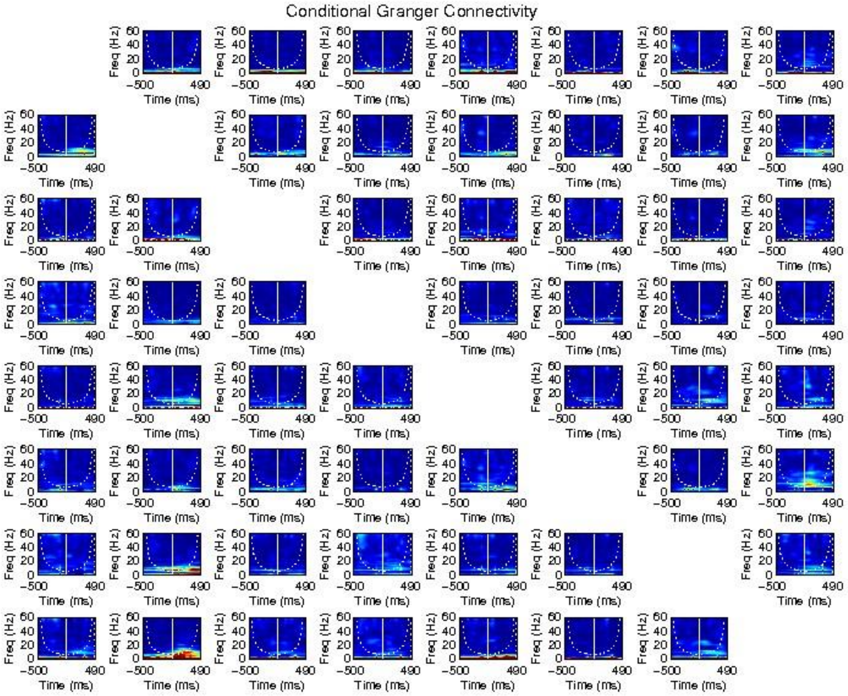
Subject 13



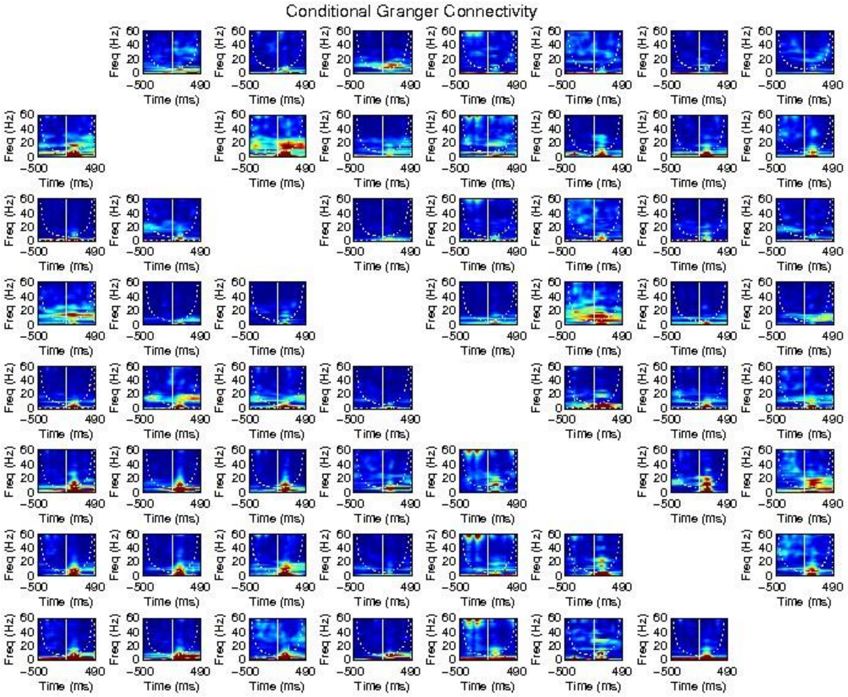
Subject 14

Figure B1: Facial Emotion Individual Subjects' Granger Spectra

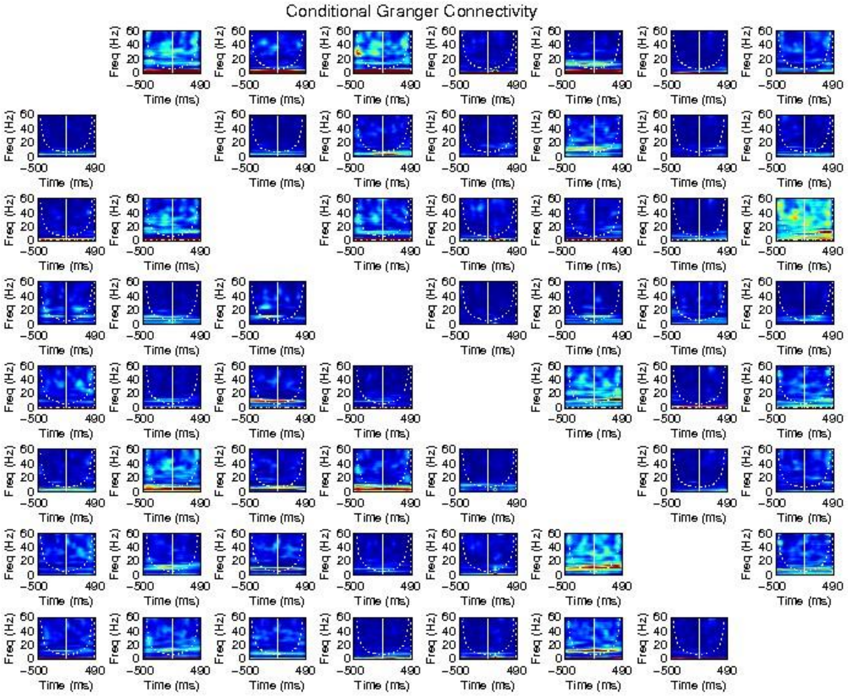
Appendix C: Facial Emotion Group Model Granger Spectra



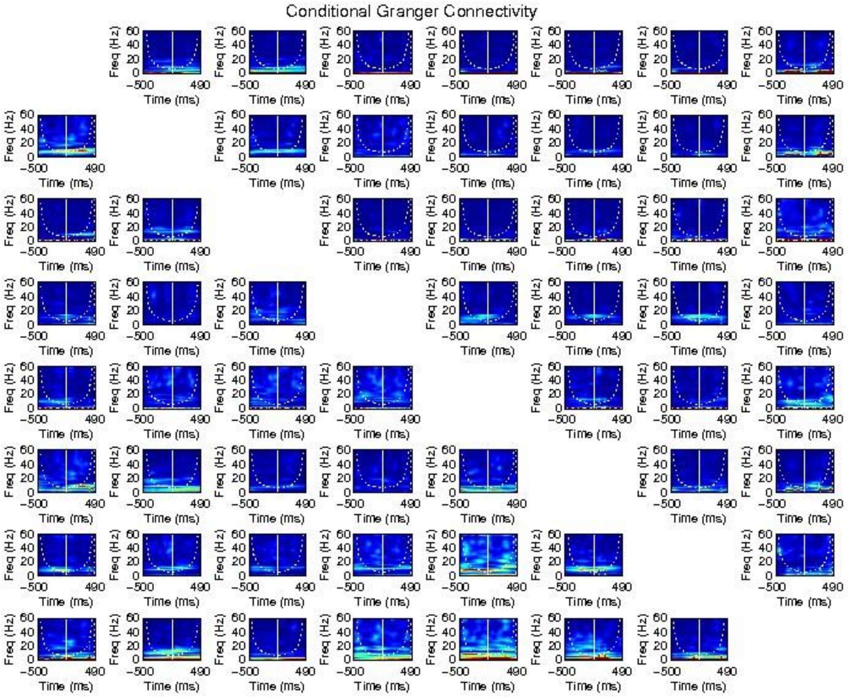
Subject 0



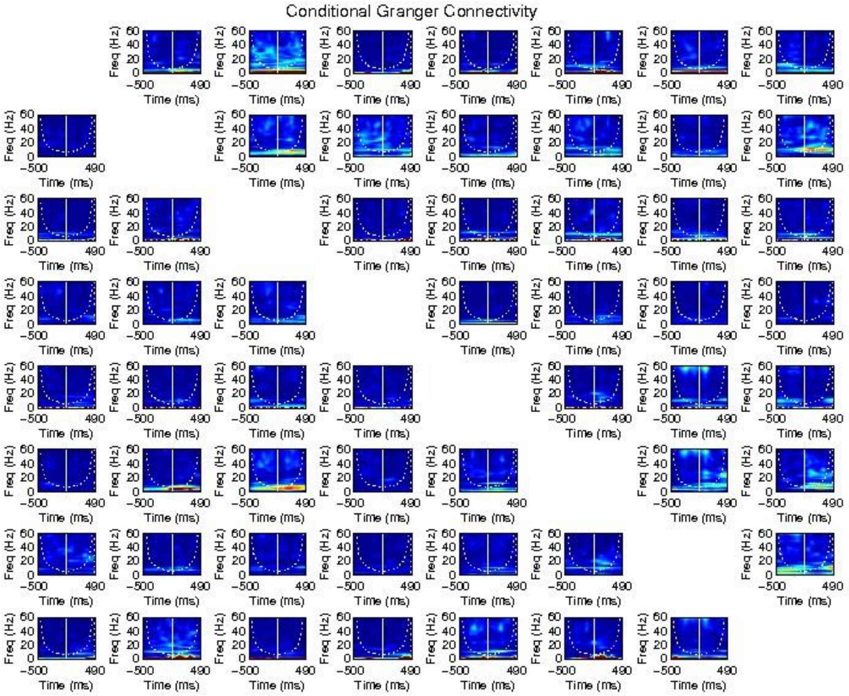
Subject 1



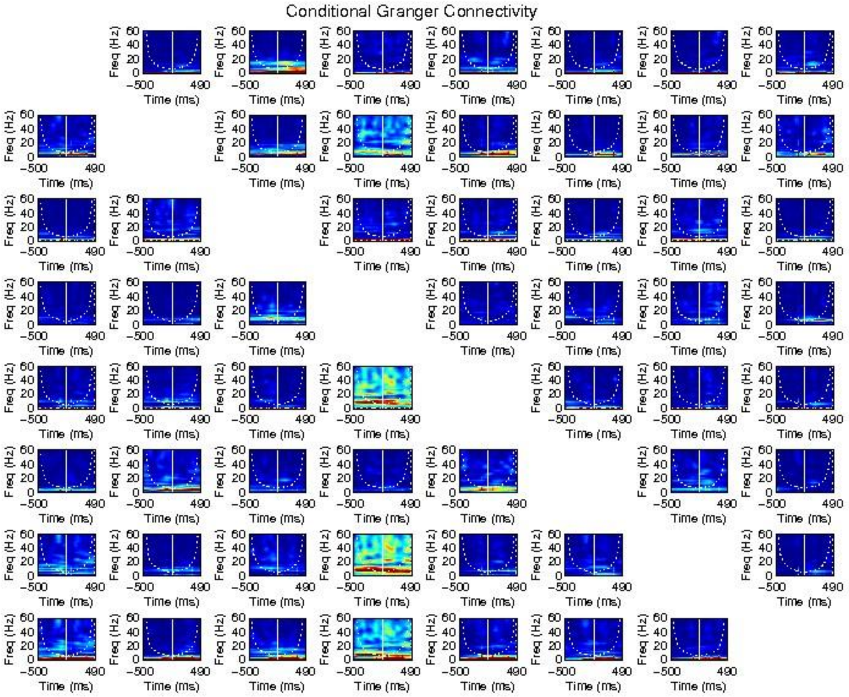
Subject 2



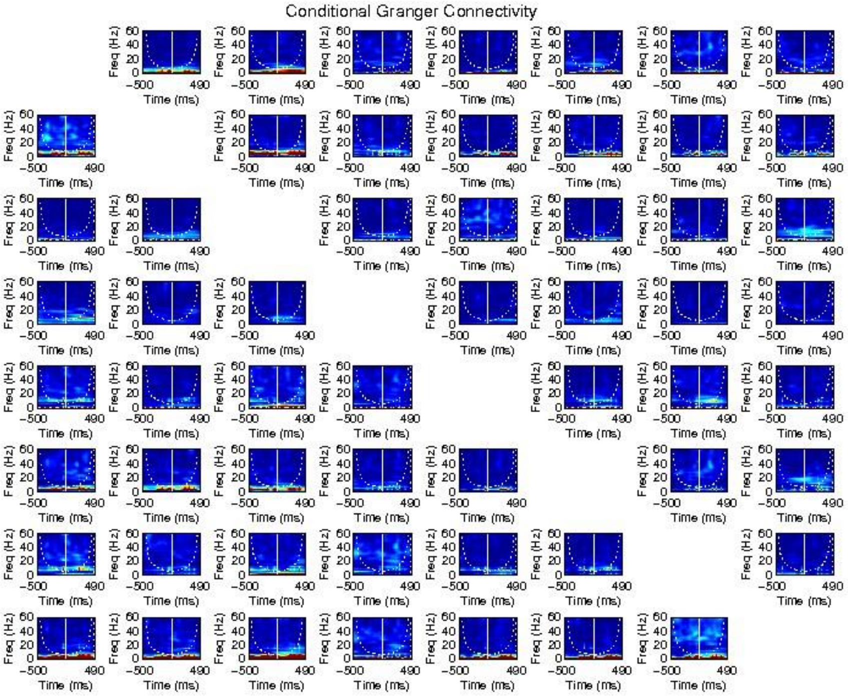
Subject 3



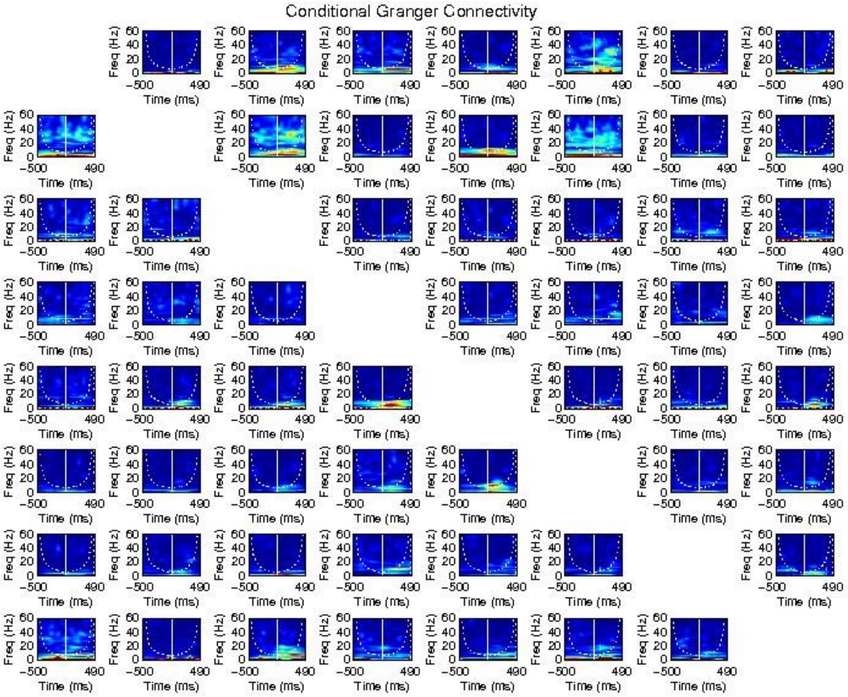
Subject 4



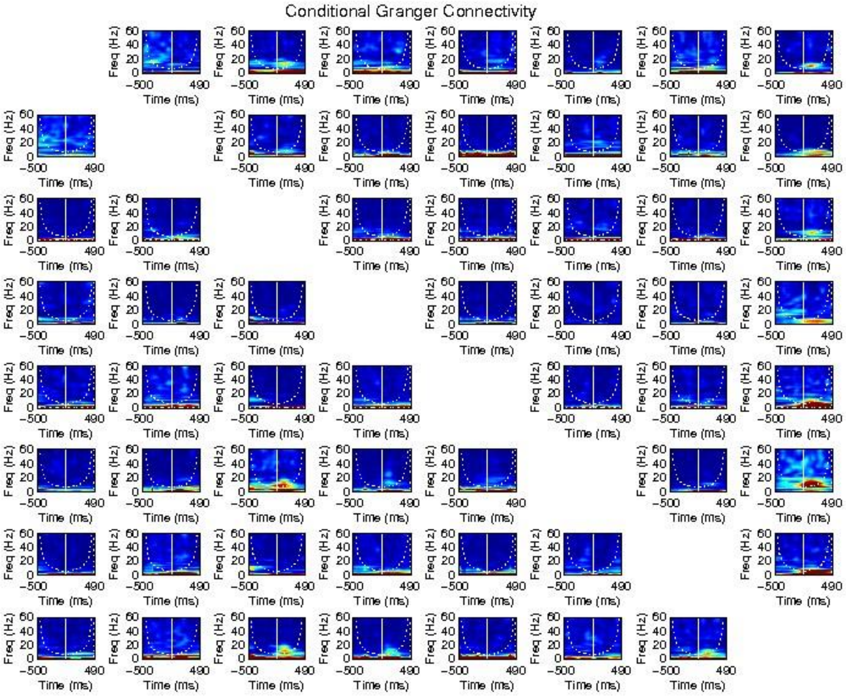
Subject 5



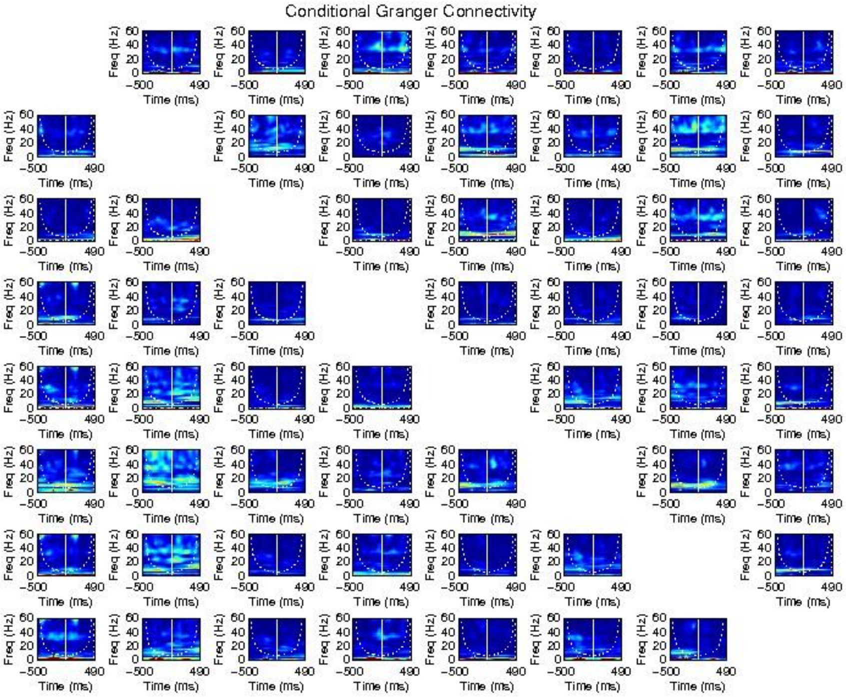
Subject 6



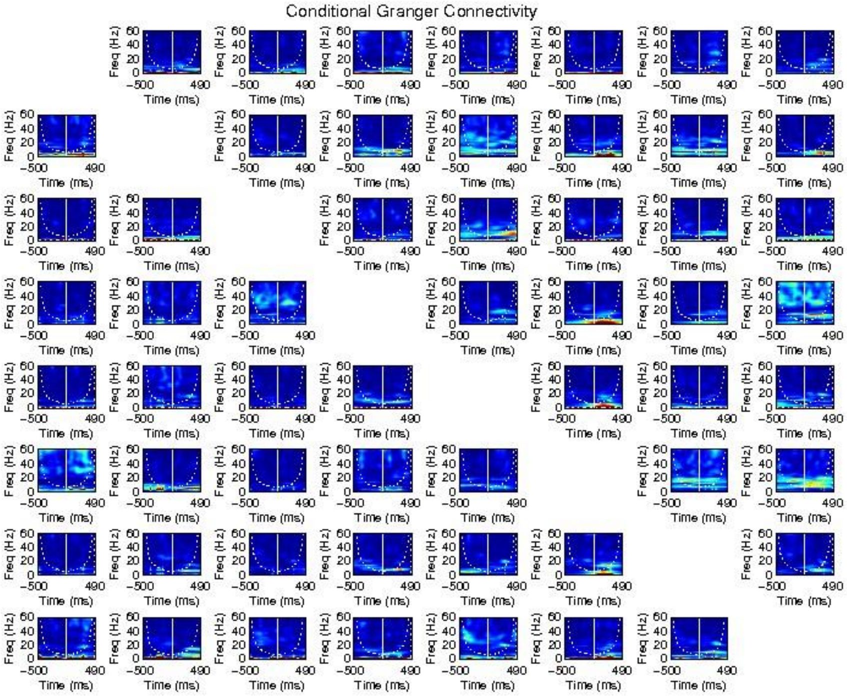
Subject 7



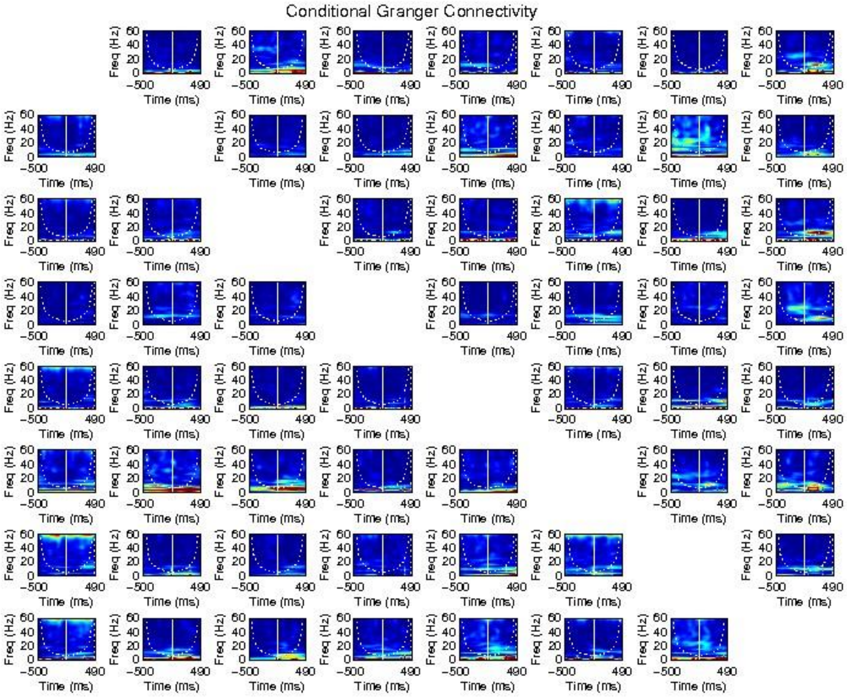
Subject 8



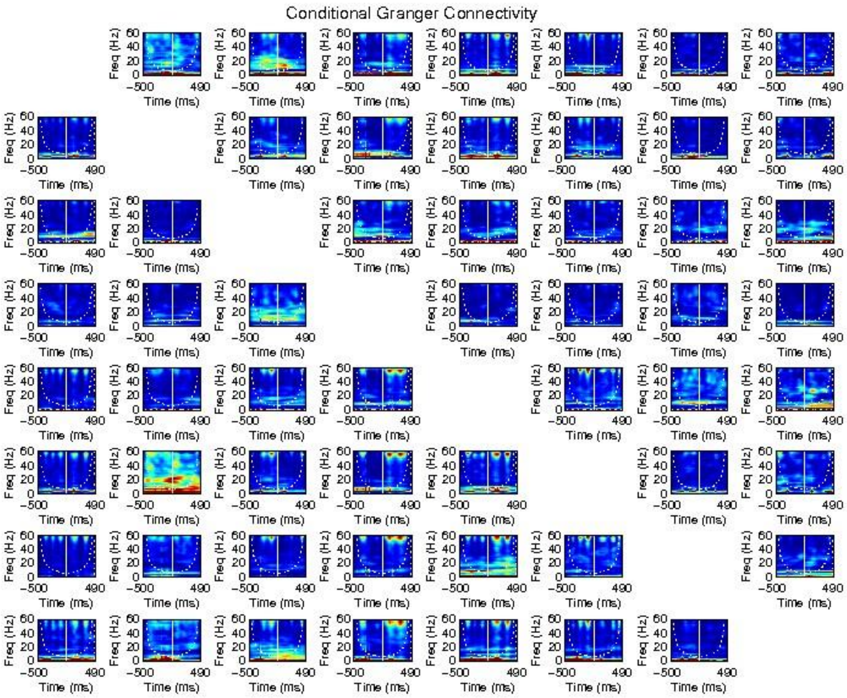
Subject 9



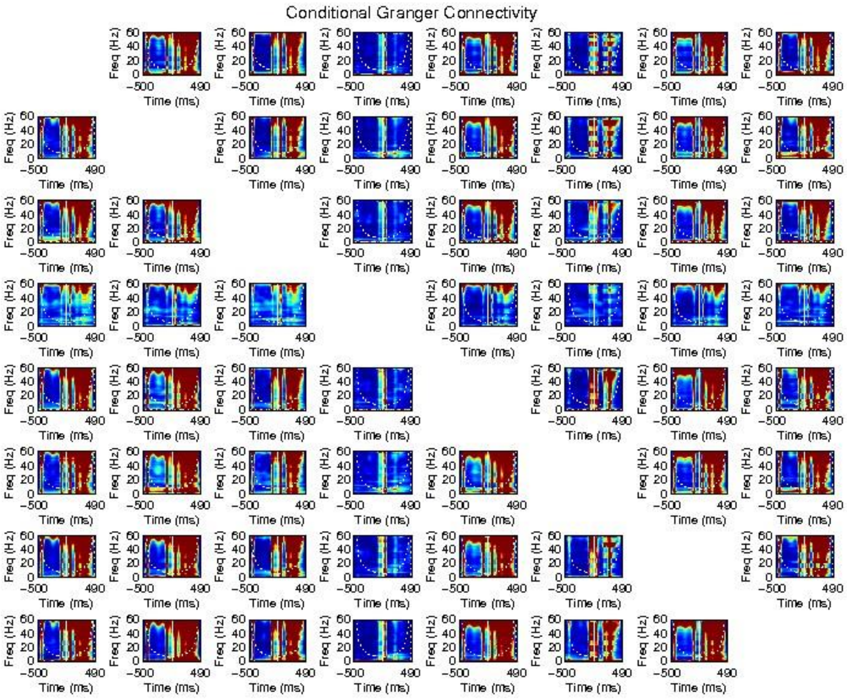
Subject 10



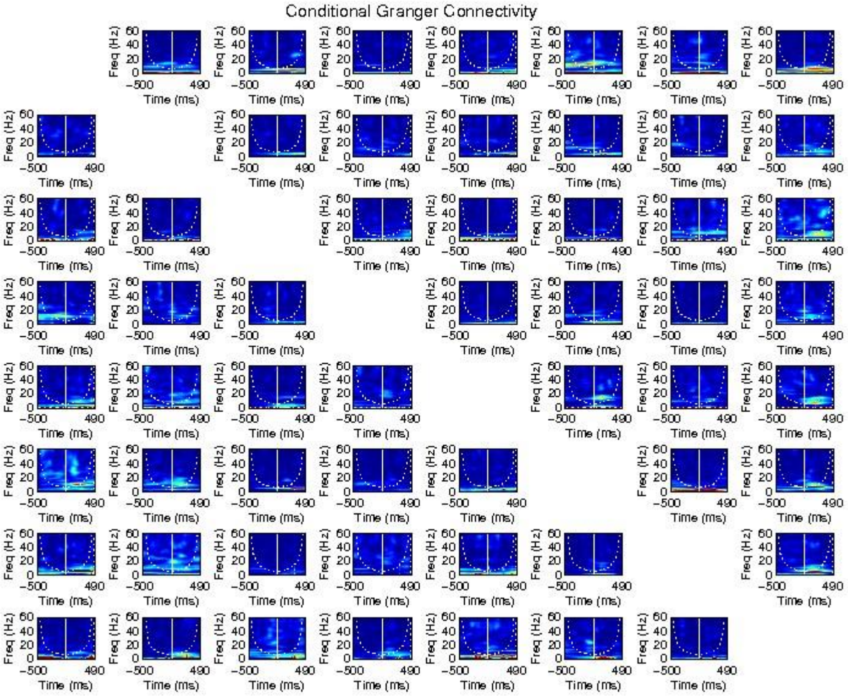
Subject 11



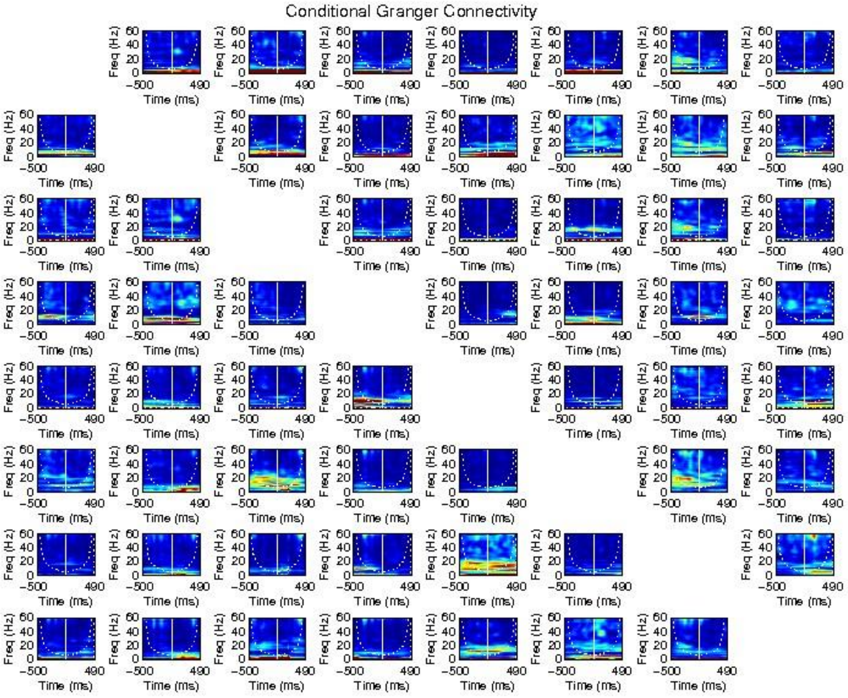
Subject 12



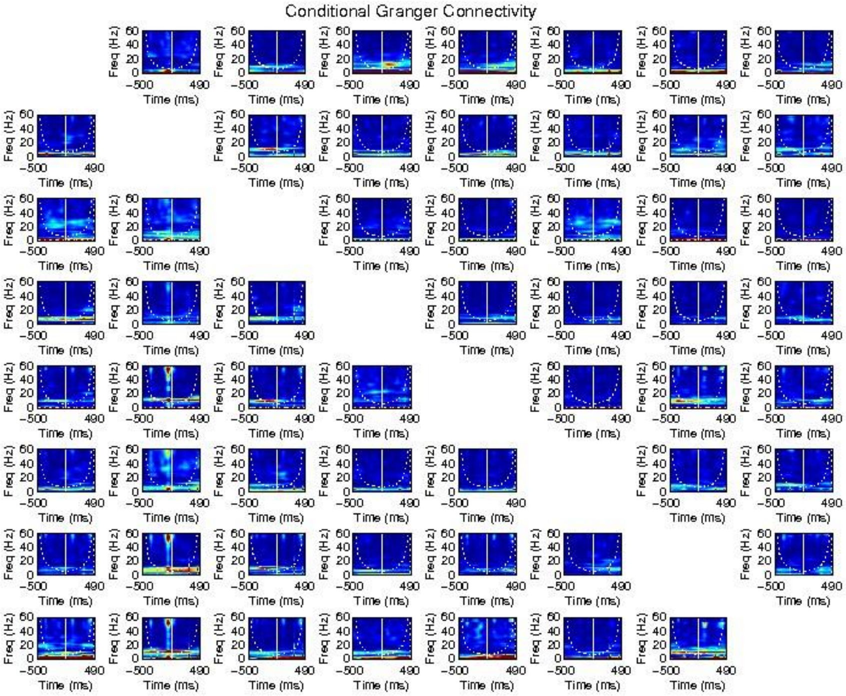
Subject 13



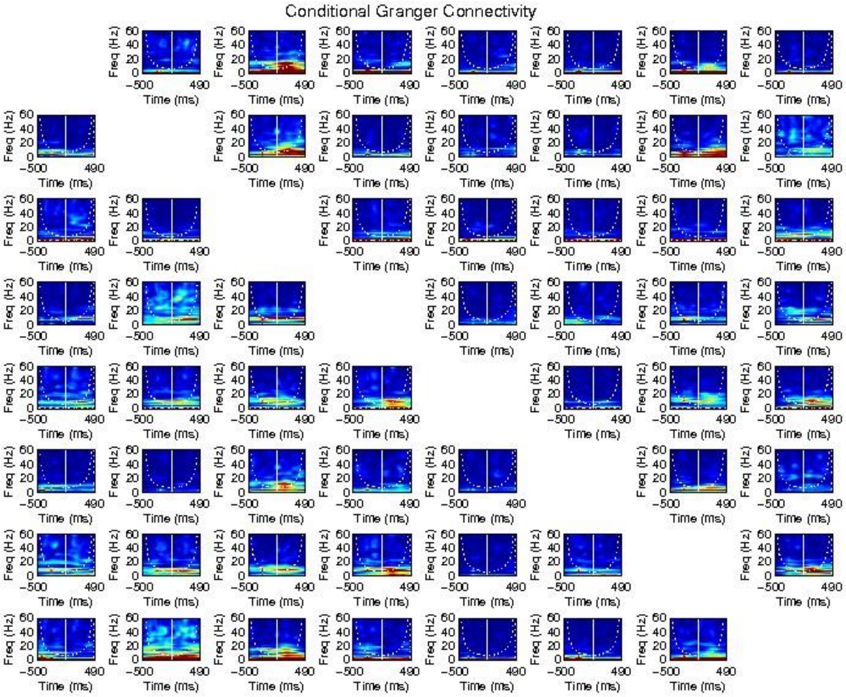
Subject 14



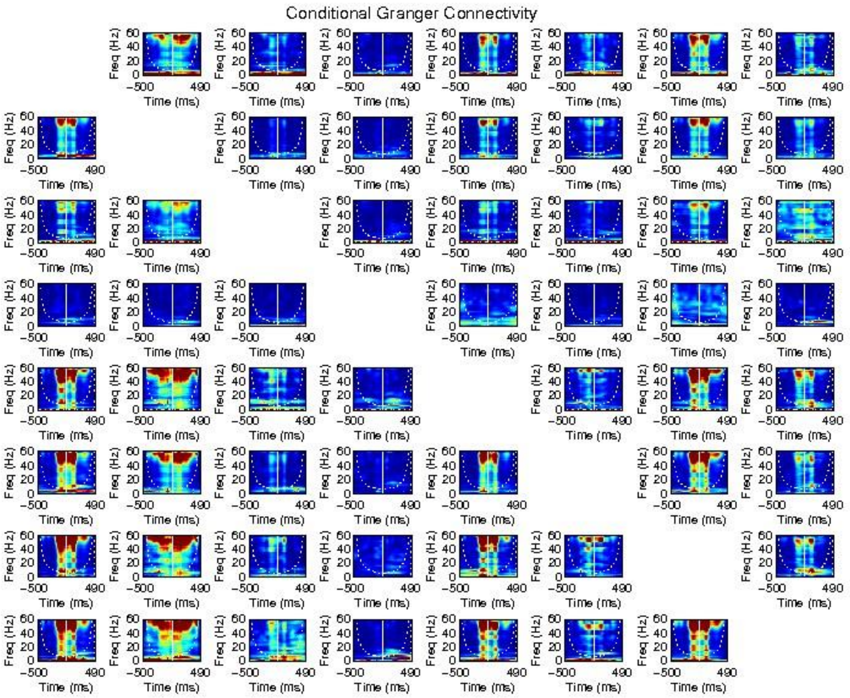
Subject 15



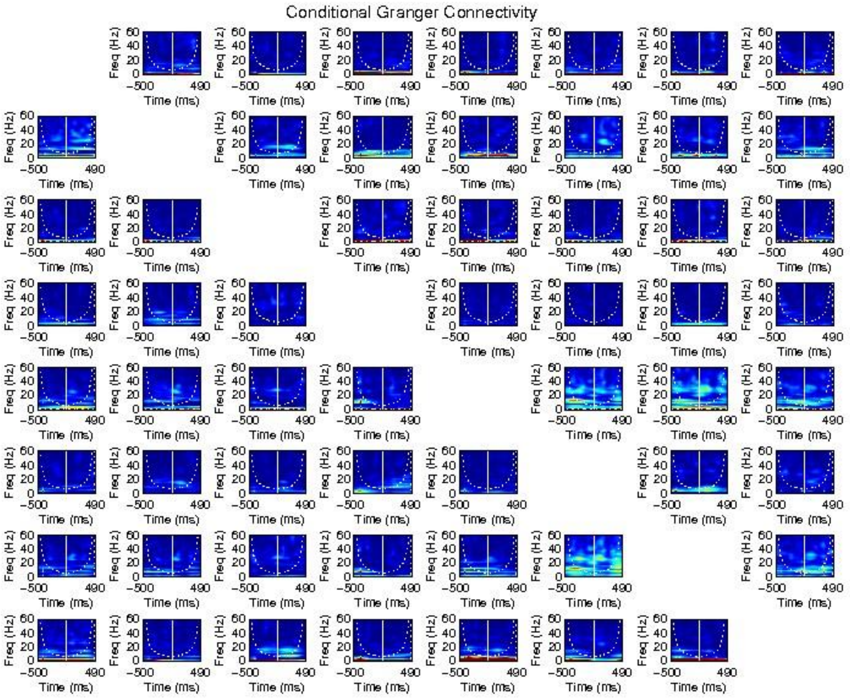
Subject 16



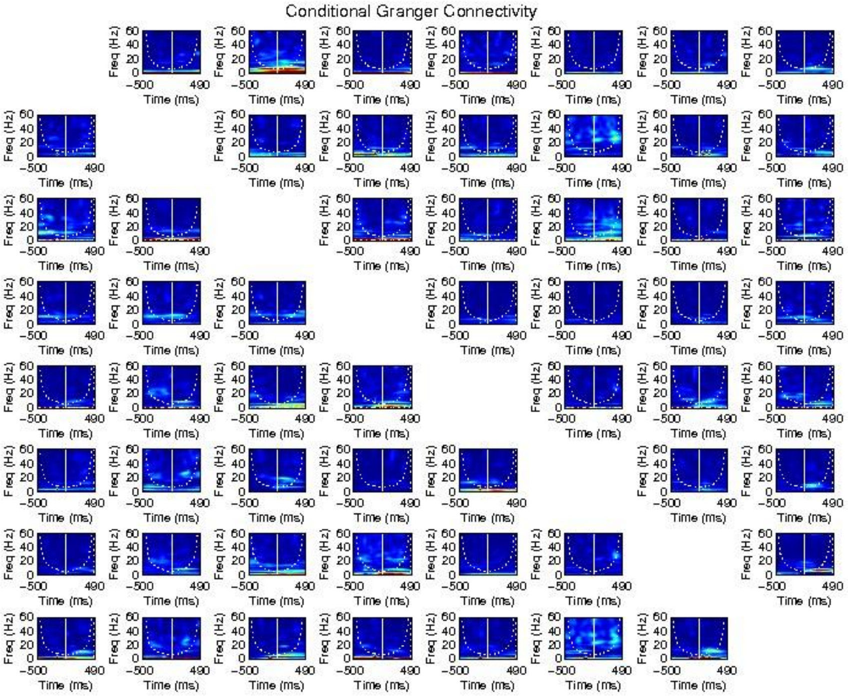
Subject 17



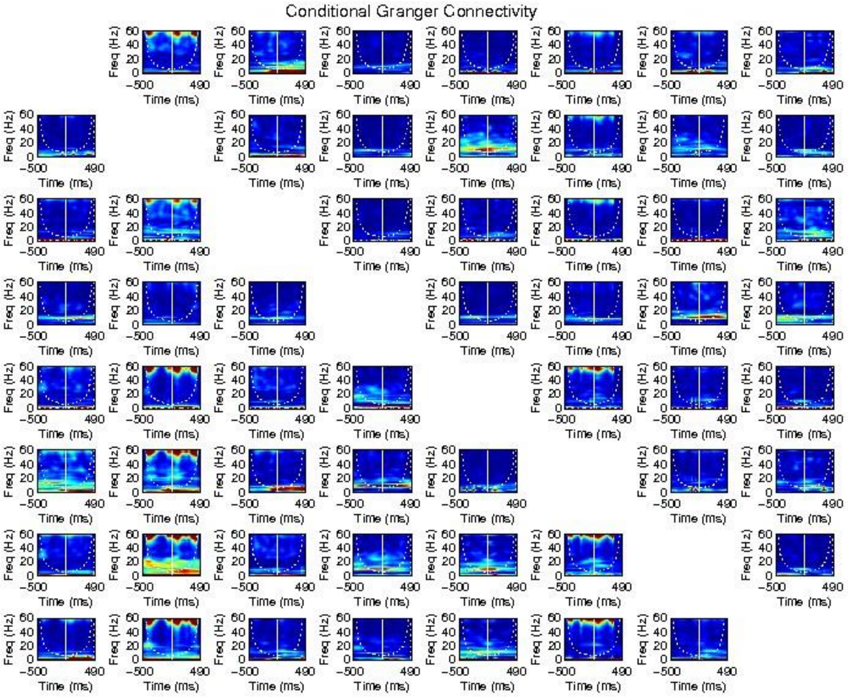
Subject 18



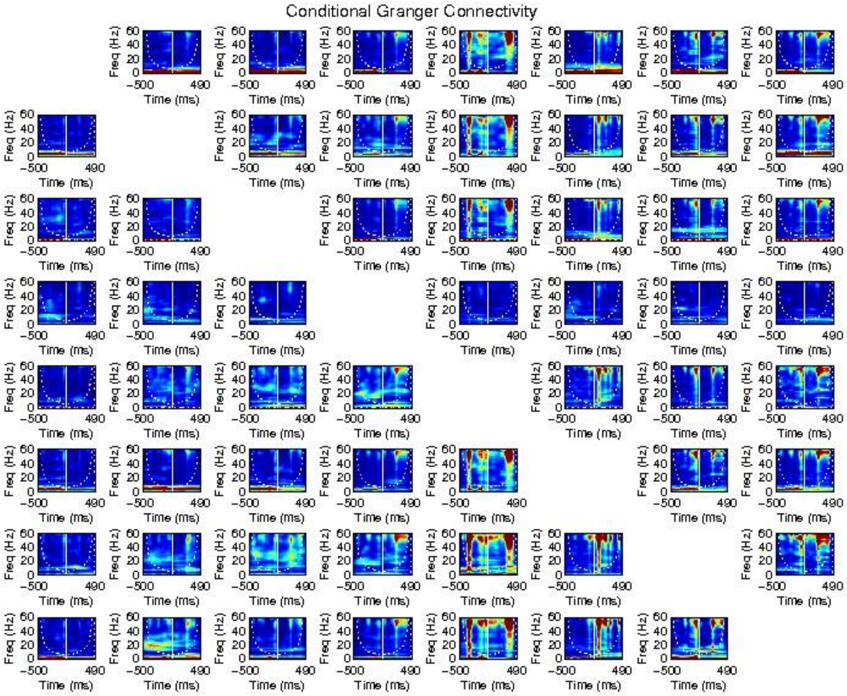
Subject 19



Subject 20



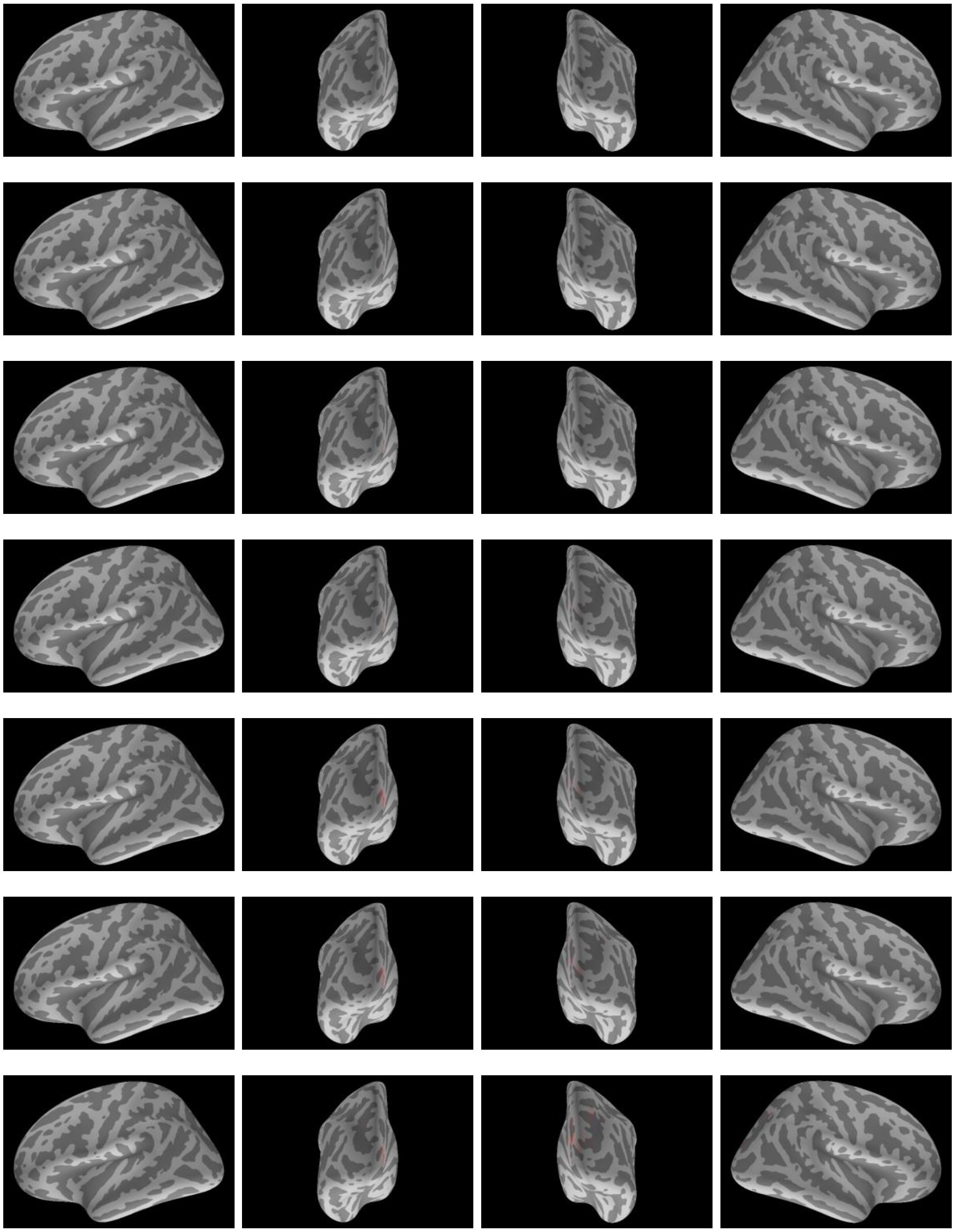
Subject 21

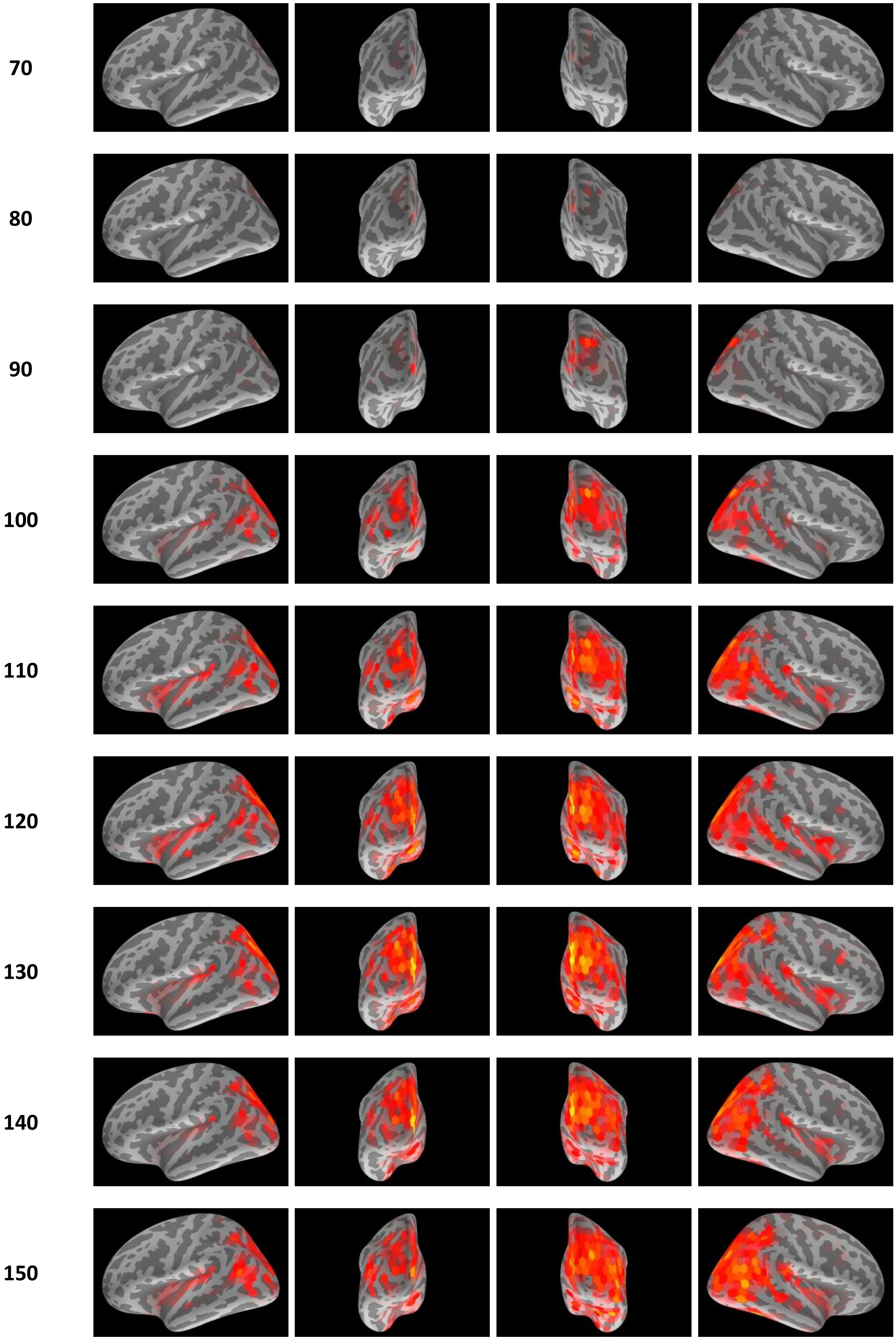


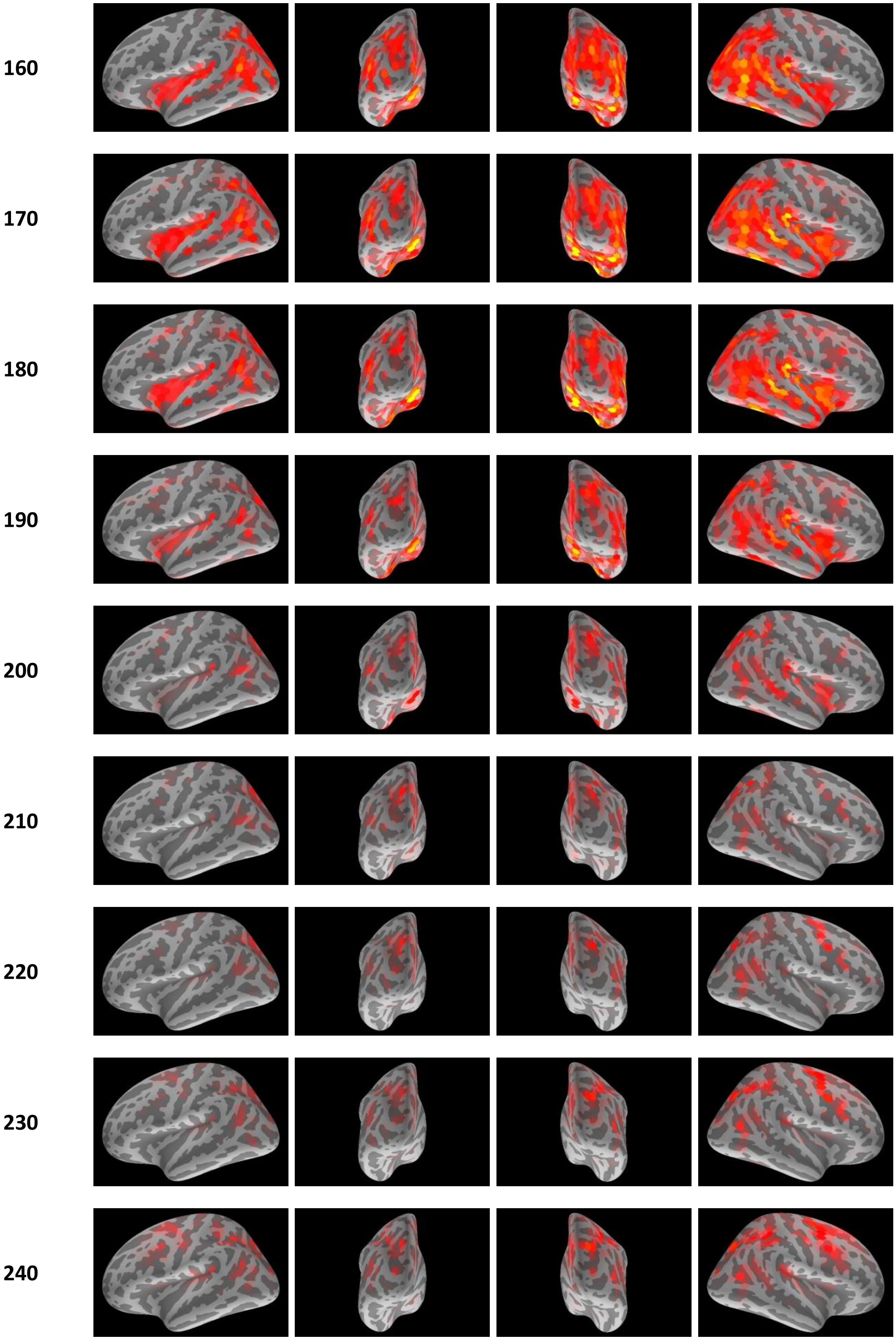
Subject 22

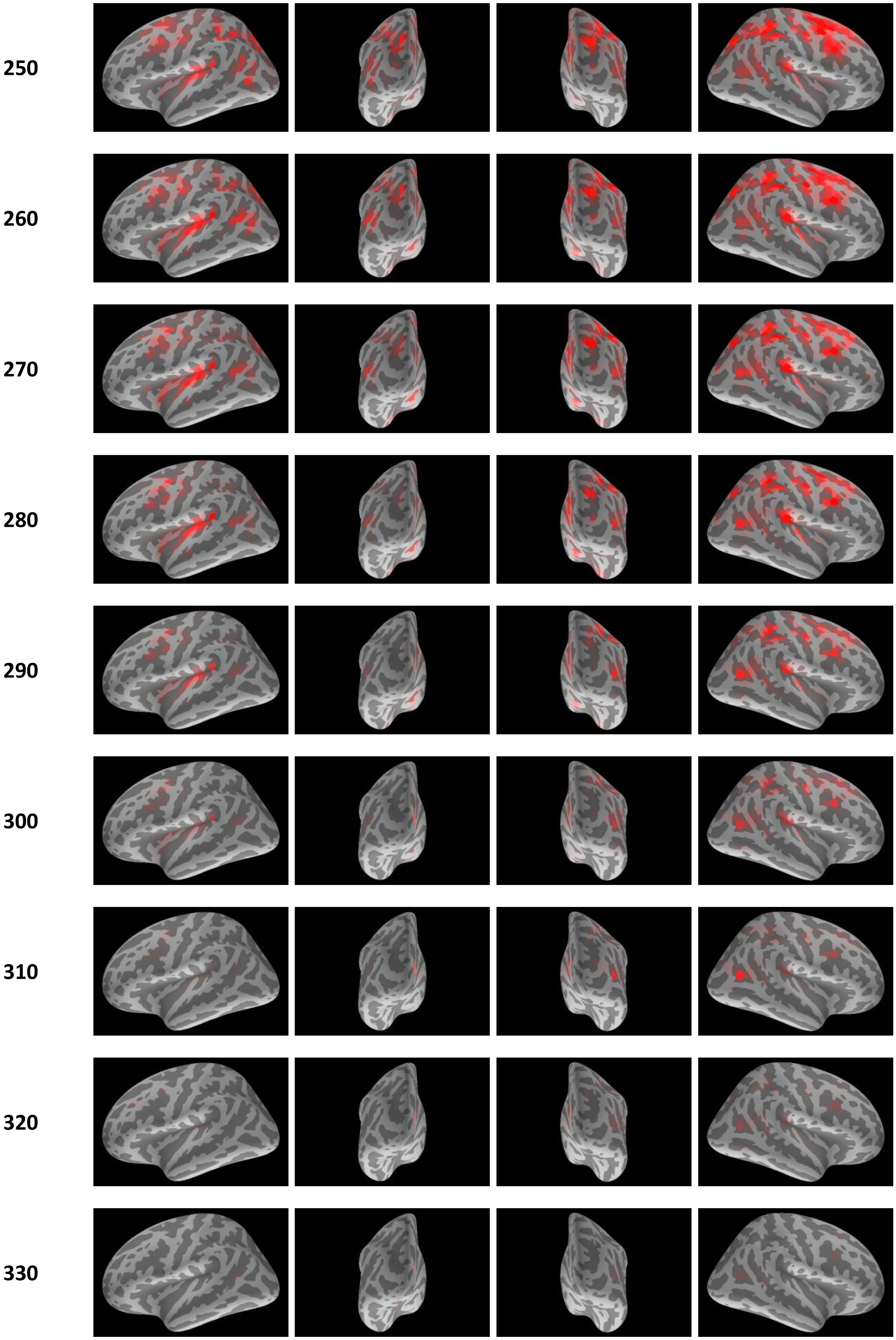
Figure C1: Facial Emotion Group Model Granger Spectra

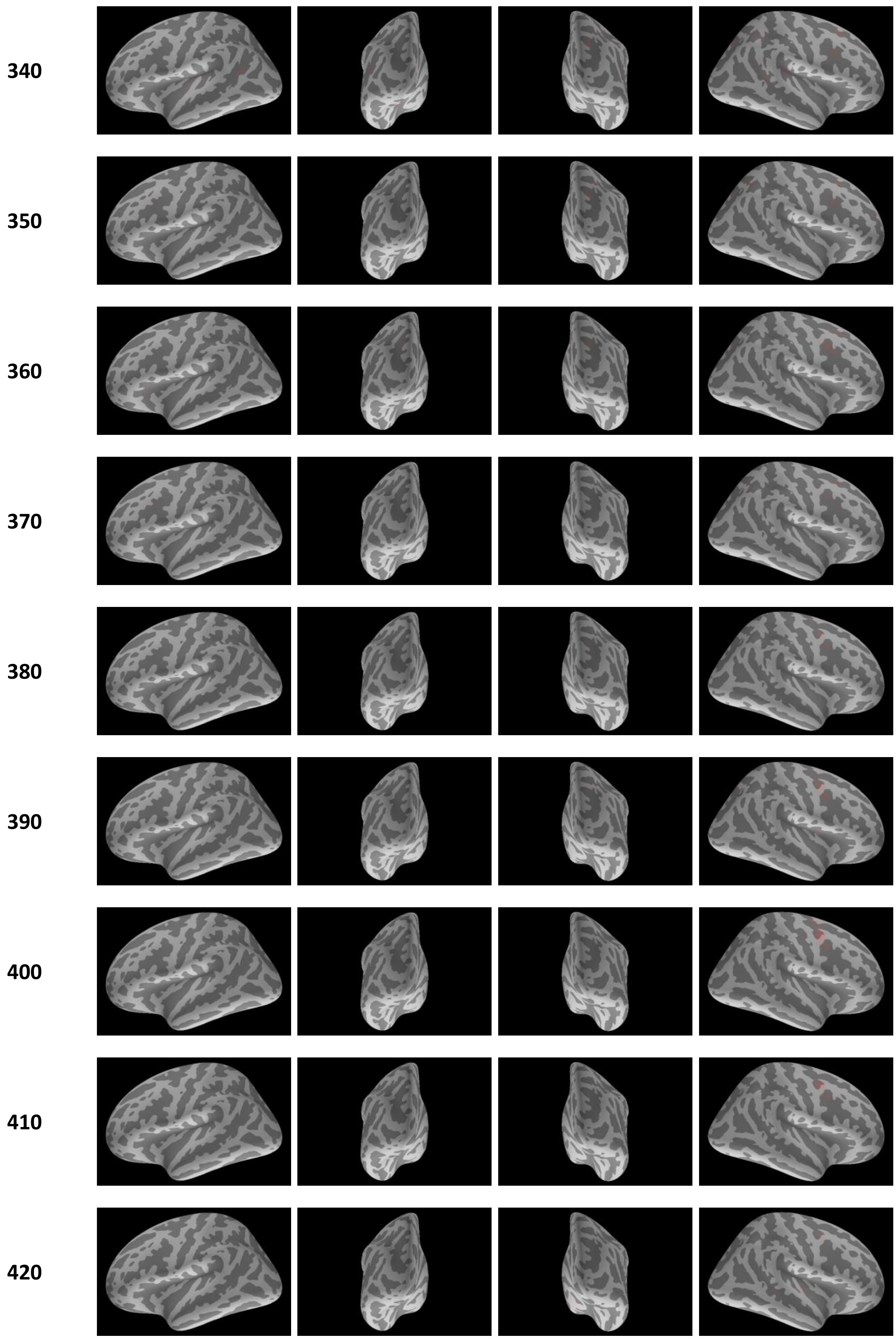
Appendix D: Facial Emotion Group L2-Norm Movie Frames




70
80
90
100
110
120
130
140
150


160
170
180
190
200
210
220
230
240


250
260
270
280
290
300
310
320
330


340
350
360
370
380
390
400
410
420

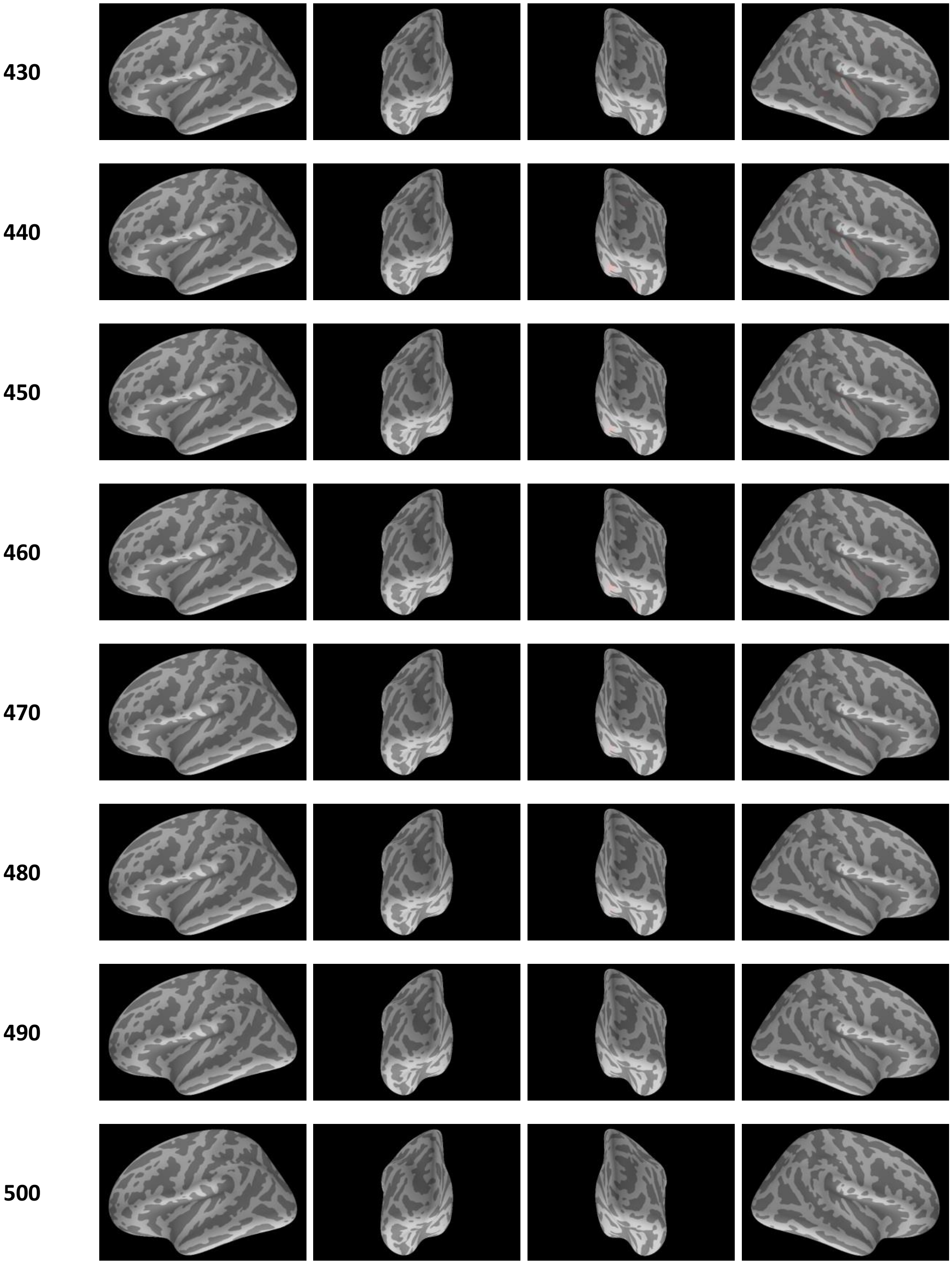


Figure D1: Facial Emotion Group L1 Frames

Appendix E: Frequency Skew in Dhamala & Rangarajan's Implementation of Wilson's Matrical Decomposition Algorithm

Dhamala and Rangarajan have generously provided a Matlab implementation of Wilson's spectral decomposition algorithm to the research community (Dhamala, Rangarajan, & Ding, 2008a; Dhamala, Rangarajan, & Ding, 2008b; Wilson, 1972; Wilson, 1978). This function is called *sfactorization_wilson* and is found in file *csd2transfer.m* of the Fieldtrip software (Oostenveld, Fries, Maris, & Schoffelen, 2011). The purpose of the Wilson algorithm is to facilitate the decomposition of the power and cross power spectra of a bivariate MVAR system by direction of influence, and a well-conceived explanation of the mechanics involved is provided by Ding et al. (Ding, Chen, & Bressler; Geweke, 1982). Unfortunately, a slight skew of peak frequency towards the lower end of the spectrum is seen when applied to the wavelet but not Fourier frequency domains.

Frequency Skew

In the bivariate MVAR system below, influence between series x_1 and x_2 exists in only one direction: from x_2 to x_1. In this special case the entirety of the cross spectra is assigned to the $x_2 \rightarrow x_1$ direction, with the opposite direction, $x_1 \rightarrow x_2$, being zero at all frequencies. This makes it possible to test the Wilson decomposition against a known outcome. Note that this is the same system Dhamala et al. (2008a; 2008b), and seen in Equations 15 and 16 above. The sampling frequency is assumed to be 200 Hz, matching the publications.

$$x_1(t) = 0.55x_1(t-1) + 0.80x_1(t-2) + 0.25x_2(t-1)$$

$$x_2(t) = 0.55x_2(t-1) + 0.80x_2(t-2)$$

This test was performed in both the Fourier and wavelet domains. In the Fourier domain there are three methods available to obtain the cross power spectral decomposition of this system:

1. Empirical deduction: Since we know this is a degenerate case, the entire measured cross power spectrum may be assigned to the direction $x_2 \rightarrow x_1$. The opposite direction, $x_1 \rightarrow x_2$, may be set to zero for all frequencies.
2. Estimation of the MVAR coefficients: When the input is either a known MVAR system as is the case here, or when the MVAR coefficients have

been estimated (Schlogl, 2006), Granger's analytic forms may be used to calculate the directional cross spectral components from the MVAR coefficients (Granger, 1969).

3. Wilson's decomposition: Despite this being a degenerate case, Wilson's decomposition algorithm may be applied.

The output from method #3 should match results from #1-2 in this degenerate case. Methods #2 and #3 should show good agreement for any MVAR system.

In the wavelet domain there is no analog to Granger's analytic formulas, so only options #1 and #3 are available. Thus in the wavelet domain the Wilson algorithm must be tested against a degenerate system such as the one used here.

The peak skew of concern in this appendix occurs in the wavelet domain but not the Fourier domain. In Figure E1 the Granger measures of influence (Cxy, Cyx) are generated using the output from the Wilson algorithm in both the Fourier (cyan, magenta) and wavelet domains (blue, red). The Fourier Cyx term shows a peak at 40 Hz (corresponding to the analytic forms, which are not show). The wavelet Cyx term shows a similar peak skewed downwards to nearly 30Hz. As expected, the Cxy terms are nearly zero at all frequencies. Cyx is also plotted as generated from the raw cross spectral power in the Fourier (magenta with dots) and wavelet (red with dots) domains. In the Fourier domain, the two peaks at 40 Hz correspond nicely, indicating agreement between the Wilson algorithm and the raw results from the degenerate system. In other words, the algorithm decomposed the cross spectra as expected in the Fourier domain. This is not the case in the wavelet domain, where decomposed results peaking at 30 Hz do not correspond to the raw cross spectrum peaking slightly above 40 Hz.

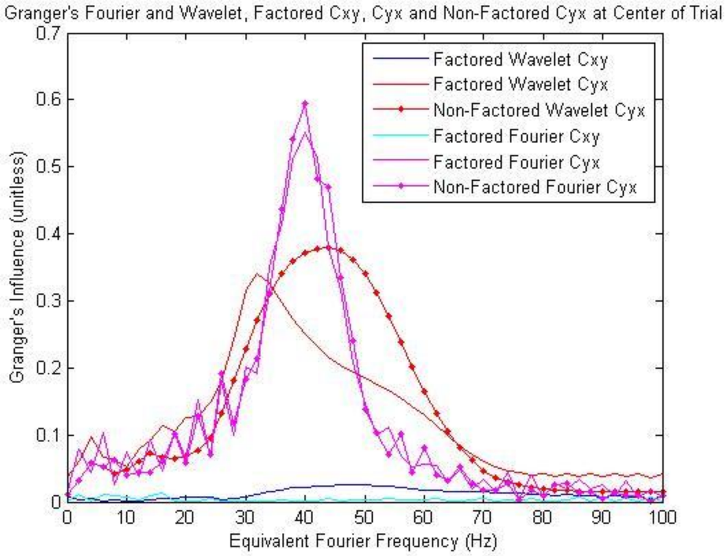


Figure E1: Granger's C_{xy}, C_{yx} in Fourier and Wavelet Domains

Comparison of the Fourier and wavelet domain values of Granger's Cxy and Cyx for the test system. Results based on the Wilson decomposition are compared against those based on the raw cross spectra (allowable since influence is known to be unidirectional) in both the Fourier and wavelet domains. The plot contains six traces representing the factored Fourier results in both directions (cyan, magenta) , the factored wavelet results in both directions (blue, red), and results based on the raw cross spectra from x2 to x1 in the Fourier and wavelet domains (red with dots, magenta with dots). In the Fourier domain the raw and factored Cyx demonstrate the same peak at 40 Hz (magenta with dots vs. magenta). In the wavelet domain the raw Cyx shows a similar peak slightly higher than 40 Hz, but the factored Cyx shows this peak skewed downwards towards 30 Hz (red with dots vs. red). This figure shows that peak frequencies based on the Wilson algorithm match the peak frequencies in a unidirectional system which does not require factorization in the Fourier but not wavelet domains. The mother wavelet used to generate this figure uses parameter w_0=6 as is commonly recommended.