UNIVERSITY OF CALIFORNIA,
IRVINE

Analysis of High-Dimensional Time Series with Applications on Brain Signals

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Statistics

by

Yuxiao Wang

Dissertation Committee:
Professor Hernando Ombao, Chair
Professor Michele Guindani
Professor Zhaoxia Yu

2017

DEDICATION

To my parents, whose love is the foundation of my life.
To Xuan, my beloved wife, whose care and support made it possible for me to complete this work.
To our forthcoming daughter, who is our most valuable gift.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ALGORITHMS

ACKNOWLEDGMENTS

This Ph.D. journey is an exciting and rewarding one which can never be completed alone by myself. May I pay tribute to those who have helped bring to this point in this special part of the dissertation.

I am very grateful to the guidance offered by my dissertation committee members, Professor Hernando Ombao, Professor Michele Guindani and Professor Zhaoxia Yu. Especially, I would like to express my deepest gratitude to my advisor Professor Hernando Ombao, for his help, encouragement and support during my entire PhD study. Without these, this dissertation would be impossible.

I would like to extend my thanks to our collaborator Dr. Chee-Ming Ting from Universiti Teknologi Malaysia, whose knowledge and insightful inputs made this dissertation as it is now.

I would also like to thank all the members in the UCI Space-Time group for their help and all the insightful discussions.

This work is supported by the Department of Statistics at UC Irvine and the Graduate Dean's Dissertation Fellowship (GDDF).

CURRICULUM VITAE

Yuxiao Wang

EDUCATION

Doctor of Philosophy in Statistics **2012-2017**
University of California, Irvine *Irvine*, CA, USA

Bachelor of Science in Applied Physics **2006-2010**
University of Science and Technology of China *Hefei*, Anhui, China

RESEARCH EXPERIENCE

Graduate Student Researcher **2012–2017**
University of California, Irvine *Irvine*, CA, USA

TEACHING EXPERIENCE

Teaching Assistant/ Grader **2012–2016**
University of California, Irvine *Irvine*, CA, USA

Instructor **2015–2017**
Data Science Initiative, University of California, Irvine *Irvine*, CA, USA

PROFESSIONAL EXPERIENCE

Research Software Engineer Intern **2016.06 - 2016.09**
Google Inc. *Mountain View*, CA, USA

Research Scientist Intern **2015.06 - 2015.09**
Amazon.com, Inc. *Seattle*, WA, USA

Software Engineer Intern **2014.06 - 2014.09**
The Mathworks, Inc. *Natick*, MA, USA

PUBLICATIONS

- Wang, Y., Ting, C. M., and Ombao, H. (2016). Modeling Effective Connectivity in High-Dimensional Cortical Source Signals. *IEEE Journal of Selected Topics in Signal Processing.*
- Wang Y, Hu, L and Ombao, H. (2016). Statistical Analysis of Electroencephagrams. In H. Ombao, M. Lindquist, W. Thompson and J. Aston (Eds). *Handbook of NeuroImaging Data Analysis.* New York: CRC Press. ISBN 9781482220971.
- Wang Y., Ting, C. M., and Ombao, H. (2016). Exploratory Multi-scale Analysis of High Dimensional Time Series. (submitted, winner of ENAR Distinguished Student Paper Award).

SOFTWARE

XHiDiTS toolbox `https://goo.gl/uXc8ei`
Matlab toolbox for Exploratory Analysis of High-Dimensional Time Series

HONORS / AWARDS

- ENAR Distinguished Student Paper Award, 2017
- Graduate Deans Dissertation Fellowship, UC Irvine 2016
- ICS Conference Presentation Grant, UC Irvine 2015 - 2017
- Chair's Fellowship, School of Information and Computer Sciences at UC Irvine 2012 - 2016
- Outstanding Undergraduate Scholarship at Univ of Sci & Tech of China (USTC) 2006 - 2010
- National Mathematics Olympiad (China) first prize (twice, in 2004 and 2005)

ABSTRACT OF THE DISSERTATION

Analysis of High-Dimensional Time Series
with Applications on Brain Signals

By

Yuxiao Wang

Doctor of Philosophy in Statistics

University of California, Irvine, 2017

Professor Hernando Ombao, Chair

Neuronal populations behave in a coordinated manner both during resting-state and while executing tasks such as learning and memory retention. One of the major challenges to analyzing brain signals such as electroencephalograms (EEGs) and functional magnetic resonance imaging (fMRI)is high dimensionality. There can be hundreds of channels in a typical EEG recording, and the number of voxels in a fMRI recording can be hundreds of thousands. We developed computationally efficient and theoretically justified tools for analyzing high dimensional brain signals. Our approach is to extract the optimal lower dimensional representations for each brain region and then characterize and estimate connectivity between regions through these factors. This approach is motivated by the fact that electroencephalograms (EEGs) from many channels within each region exhibit a high degree of multicollinearity and synchrony thereby suggesting that it would be sensible to extract summary factors for each region. We focus on two types of linear filters. These methods were compared through simulations under different conditions and the results provide insights on advantages and limitations of each. We also performed exploratory analysis of resting state EEG data and fMRI data. The spectral properties of the factors were estimated and connectivity between regions via the factors using coherence measures were computed. We implemented these methods in a Matlab toolbox XHiDiTS. The toolbox was utilized to investigate consistency

of these factors across all epochs during the entire resting-state period. In order to quantify the effective connectivity among sources in a densely voxelated (high-dimensional) cortical surface, we developed the source-space factor vector autoregressive (VAR) model. The first step in our procedure is to estimate cortical activity from multichannel electroencephalograms (EEG) using anatomically constrained brain imaging methods. Following parcellation of the cortical surface into disjoint regions of interest (ROIs), latent factors within each ROI are computed using principal component analysis. These factors are ROI specific low-rank approximations (or representations) which allow for efficient estimation of connectivity in the high-dimensional cortical source space. The second step is to model effective connectivity between ROIs by fitting a VAR model jointly on all the latent processes. The different cortical sources within a ROI may share common factors as each source is a mixture of these VAR factors. From this commonality we derive the connectivity between the sources. Measures of cortical connectivity, in particular partial directed coherence (PDC), are formulated using the VAR parameters. We illustrate the proposed model to investigate connectivity and interactions between cortical ROIs during resting state.

Chapter 1

Introduction

1.1 Background on Electroencephalography (EEG)

Electroencephalography (EEG) is the recordings of the electrical signal on the scalp. It measures the fluctuation of voltages due to the neuron activities (Niedermeyer and da Silva, 2005). The first EEG recording on human was performed by Hans Berger in 1924 (Berger, 1929; Haas, 2003). In general, EEG waveforms are classified using their amplitude, shape and frequency. One classification of EEG waveforms is based on the frequency range where each frequency band of EEG is closely related to brain activity. The frequency bands are delta (less than 4 Hz), theta (4-8 Hz), beta (16-32 Hz) and Gamma (32-50 Hz) bands. In cognitive studies, EEGs have been widely used because of its advantages of high temporal resolution (in milli-second), non-invasiveness and low cost. It is also of interest to study the properties of brain activities from EEG data. The remainder of this introduction chapter overviews (1.) modeling EEG sources; (2.) the approaches for EEG source reconstruction; and (3.) brain connectivity measures. Part of this chapter has already been published in the book chapter Wang et al. (2016a).

1.1.1 EEG Sources

EEG signals are recordings on the scalp which means that EEGs are not direct measurements of neuronal activities on cortex. In order to model the cortical source activity from EEG recordings, two types of models have been developed. One class of models depends on the physical head forward model which assumes the EEG sources are dipole currents that are located within brain region. The other class of models assumes that the sources represent independent brain networks, which might be either localized or distributed.

Dipole Source Model

According to (Srinivasan and Deng, 2012), the potential at location r in the brain or on the scalp surface can be expressed as the integral of the contributions of sources

$$X(r,t) = \int_B G_H(r,r')P(r',t)dV(r') \tag{1.1.1}$$

where $P(r',t)$ represents the dipole moment per unit volume at location r', where r' ranges over the whole brain region B; $G_H(r,r')$ is Green's function which depends on the properties of the volume conductor and the locations of source r' and measurement location r. In order to formulate the EEG inverse problem in a fundamental point of view, according to (Nunez and Srinivasan, 2006a), the entire brain volume can be parceled into N voxels, each of volume ΔV, each having a strength $p_n(r_n,t) = P(r_n,t)\Delta V$. The potential given by Equation 1.1.1 can be written as a finite sum over contributions from N voxels:

$$X(r_k,t) = \sum_{n=1}^{N} G_n(r_k,r_n)p_n(r_n,t)\ . \tag{1.1.2}$$

According to (Baillet et al., 2001a), the algebraic formulation of EEG measurement at location r can be written as $\mathbf{X}(r) = \sum_i a(r,r_{q_i},\Theta_{q_i})q_i$, where $a(r,r_q,\Theta_q)$ is formed as the

solution to electric forward problem for a dipole with unit amplitude and orientation Θ. For the simultaneous activation of N dipoles located at r_{q_n}, the EEG recording for M locations can be written as $\mathbf{X}(t) = \mathbb{A}\mathbf{Z}(t)$, where $\mathbf{X}(t) = [X(r_1, t), ..., X(r_M, t)]'$ is EEG measurement for M locations at time t; $A_{mn} = a(r_m, r_{q_n}\Theta_{q_n})$ is the gain matrix relating the dipoles to the sensors, and $\mathbf{Z}(t) = [q_1(t), ..., q_N(t)]'$ is the source amplitudes. Each column of $\mathbb{A}$ is the forward field of the current dipole, sampled as M discrete locations of the EEG sensors.

Independent Source Model

Section 1.1.1 models the EEG signals as the potential generated by dipole currents within brain regions due to neuronal activities, where the sources might be correlated. An alternative way of modeling EEG signals is to treat it to be the output of a number of statistically independent potential-generating systems. The systems are spatially fixed but can be either spatially restricted or widely distributed, according to (Makeig et al., 1996a). The independent systems can be interpreted as independent brain activity networks. Figure 1.1 shows

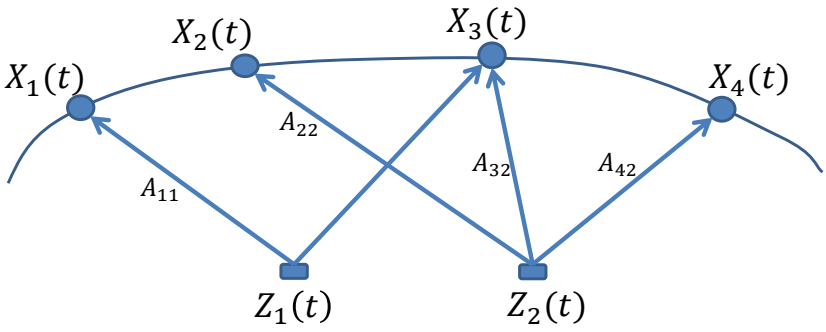


Figure 1.1: Graphical representation of latent source model.
The directions of the arrows represent dependence relationship.

a graphical illustration of the mixing structure for the signals of two independent sources (Z_1 and Z_2) and four electrodes (X_1, ..., X_4). The element A_{mn} in mixing matrix $\mathbb{A}$ can be interpreted as the loading of n-th source signal Z_n on the EEG signal at m-th electrode

(X_m). If $A_{mn} = 0$, for some m and n, we conclude that $Z_n(t)$ has no contribution to EEG signal $X_m(t)$, since the loading for source $Z_n(t)$ is zero.

1.1.2 Formulation of EEG Source Models

According to Section 1.1.1 and 1.1.1, a general model of EEG data can be written as

$$\mathbf{X}(t) = \mathbb{A}\mathbf{Z}(t) + \mathbf{N}(t) \tag{1.1.3}$$

where $\mathbf{X}(t) \in \mathbb{R}^M$ is the recorded EEG signal for M electrodes at time t; $\mathbf{Z}(t) \in \mathbb{R}^N$ represents the activity of N source signal at time t due to brain activities and is not directly observed; $\mathbb{A} \in \mathbb{R}^{M\times N}$ is a matrix representing instantaneous source mixing. The noise term $\mathbf{N}(t) \in \mathbb{R}^M$ represents either measurement error or machine error at time t.

Let $\mathbf{X} = [\mathbf{X}(1), \ldots, \mathbf{X}(T)]$, $\mathbf{Z} = [\mathbf{Z}(1), \ldots, \mathbf{Z}(T)]$, and $\mathbf{N} = [\mathbf{N}(1), \ldots, \mathbf{N}(T)]$, then we have $\mathbf{X} \in \mathbb{R}^{M\times T}$, $\mathbf{Z} \in \mathbb{R}^{N\times T}$ and $\mathbf{N} \in \mathbb{R}^{N\times T}$. The model in Equation 1.1.3 can be written in the equivalent form $\mathbf{X} = \mathbb{A}\mathbf{Z} + \mathbf{N}$.

1.1.3 Inverse Source Reconstruction

According to (Baillet et al., 2001a), the lead field matrix $\mathbb{A}$ and the source activity $\mathbf{Z}(t)$ can be estimated based on the physical modeling of the head. The lead field matrix $\mathbb{A}$ can be estimated from a spherical head model or a realistic head model. The solution for spherical head model can be computed analytically (Rush and Driscoll, 1969) however the solution for a realistic head model can only be computed numerically. The most popular approaches to estimating the lead field matrix $\mathbb{A}$ are the BEM (boundary element method (Hamalainen and Sarvas, 1989)) and FEM (finite element method (Yan et al., 1991)). The only thing unknown in inverse source reconstruction is the dipole activity q_n, $n = 1, 2, ..., N$. In general, there

are two approaches to estimation of EEG sources: parametric and imaging methods.

Parametric Methods

Parametric methods typically assume the sources can be characterized by a few equivalent current dipoles sources with unknown locations. The dipole moments are estimated using non-linear numerical methods including simplex search, genetic algorithm, and simulated annealing (Uutela et al., 1998). In the presence of noise, the forward model at time can be written as

$$\mathbf{X}(t) = \mathbb{A}\mathbf{Z}(t) + \mathbf{N}(t) \tag{1.1.4}$$

where $\mathbf{X}(t) = [X_1(t), \ldots, X_M(t)]'$ is EEG measurement at M locations $\{r_1, \ldots, r_M\}$; $A_{mn} = a(r_m, r_{q_n}\Theta_m)$ is the gain matrix relating the dipoles to the sensors, $\mathbf{Z}(t) = [q_1(t), \ldots, q_N(t)]'$ is the source amplitudes, $\mathbf{N}(t) \in \mathbb{R}^M$ is the measurement and machine noise.
Let $\mathbf{X} = [\mathbf{X}(1), \ldots, \mathbf{X}(T)]$, $\mathbf{Z} = [\mathbf{Z}(1), \ldots, \mathbf{Z}(T)]$, and $\mathbf{N} = [\mathbf{N}(1), \ldots, \mathbf{N}(T)]$, the model in Equation 1.1.4 can be written in the following equivalent form $\mathbf{X} = \mathbb{A}\mathbf{Z} + \mathbf{N}$. The goal of dipole fits is to determine best set $\{r_i, \Theta_i\}$ of number N sources and model parameters, and sources $\mathbf{Z}$ that best describe the data. Least square estimation (LSE) is one of the approaches. The LSE approach minimized the cost function in the form

$$J_{LS}(r_{q_i}, \Theta_i, \mathbf{Z}) = \|\mathbf{X} - A(\{r_{q_i}, \Theta_i\})\mathbf{Z}\|_F^2 \tag{1.1.5}$$

where the the subscript F indicates the Frobenius norm, that is

$$\|\mathbf{X} - A(\{r_{q_i}, \Theta_i\})\mathbf{Z}\|_F^2 = \mathrm{Tr}[(\mathbf{X} - A(\{r_{q_i}, \Theta_i\})\mathbf{Z})(\mathbf{X} - A(\{r_{q_i}, \Theta_i\})\mathbf{Z})^T]$$

The cost function 1.1.5 can be minimized by exhaustive scanning through the whole solution space. However it is too computational expensive, even for small values of N. Some non-linear optimization based on directed search algorithms have been developed, which are more efficient than the exhaustive searching algorithm. The non-linear optimization algorithms also suffer from the problem of easily being trapped in local optima.

One disadvantage of the dipole fits model is that the correct number of dipole sources N is unknown in priori and it has to be estimated. Approaches to estimating the number of sources include stepwise selection, multiple signal classification (MUSIC), principal component analysis (PCA), independent component analysis (ICA), second order blind identification (SOBI).

Imaging Methods

Imaging methods assume that the primary sources are intracellular currents in dendritic trunks of the pyramidal neuron in cerebral cortex. (Teplan, 2002; Baillet et al., 2001a). It performs inference on dipole moments at a fixed set of locations within the brain and therefore the inverse problem in this case is linear. Number of sensors is in the order 10^2, while number of unknowns source parameters are in the order of 10^4.

Distributed source models reconstruct the electric activity at each point of the grid of solution space (3D for brain activity, 2D for cortical activity). It makes assumption on the number of dipoles in the brain. It overcomes the problem with the diople model described in Section 1.1.3 that the exact number of dipoles could not been determined a priori. However, in the distributed source models, the number of source points N is much greater than the number of EEG electrodes M, which makes the source inversion problem ill-conditioned. In order to solve the ill-conditioned inversion problem, regularization must be applied. For example, constraints that are based on biophysical knowledge or other imaging techniques. The basic

assumptions for distributed source models are: (1.) Each point can be treated as a dipole with fixed location, the only that varies is magnitude and orientation and (2.) the observed EEG signals are linearly related to the source magnitude through a mixing matrix A.

In general, the approach for source imaging method minimizes a cost function that combines both the estimation error and a penalty term

$$\widehat{\mathbf{Z}} = \arg\min_{\mathbf{Z}} \|\mathbf{X} - \mathbb{A}\mathbf{Z}\|_Q^2 + f(\mathbf{Z}) \tag{1.1.6}$$

where the Q-norm in equation 1.1.6 is defined as follows

$$\| \mathbf{X} - \mathbb{A}\mathbf{Z} \|_Q^2 = \mathrm{Tr}[(\mathbf{X} - \mathbb{A}\mathbf{Z})^T Q^{-1} (\mathbf{X} - \mathbb{A}\mathbf{Z})],$$

with $Q = \mathbb{E}[\mathbf{N}(t)\mathbf{N}(t)^T]$ representing the variance-covariance matrix of the noise term $\mathbf{N}(t)$. The penalty term $f(\mathbf{Z})$ in equation 1.1.6 represents the a priori assumptions. Note that if $f(\mathbf{Z})$ can be written as $f(\mathbf{Z}) = \sum_{t=1}^{T} f_t(\mathbf{Z}(t))$, then Equation 1.1.6 can be reformulated to

$$\widehat{\mathbf{Z}} = \arg\min_{\mathbf{Z}} \sum_{t=1}^{T} \Big[(\mathbf{X}(t) - \mathbb{A}\mathbf{Z}(t))^T Q^{-1} (\mathbf{X}(t) - \mathbb{A}\mathbf{Z}(t)) + f_t(\mathbf{Z}(t)) \Big] \tag{1.1.7}$$

Therefore for any given time t, the estimator of $\mathbf{Z}(t)$ can be constructed using

$$\widehat{\mathbf{Z}}(t) = \arg\min_{\mathbf{Z}(t)} \Big[(\mathbf{X}(t) - \mathbb{A}\mathbf{Z}(t))^T Q^{-1} (\mathbf{X}(t) - \mathbb{A}\mathbf{Z}(t)) + f_t(\mathbf{Z}(t)) \Big] \tag{1.1.8}$$

Based on different assumptions on regularization condition, a few approaches have been developed for cortical source imaging:

- Minimum norm (MN) method

  The minimum norm estimator (Hämäläinen and Ilmoniemi, 1994a) minimizes the ob-

jective function given by

$$\| \mathbf{X} - \mathbb{A}\mathbf{Z}\|_F^2 + \lambda\|\mathbf{Z}\|_F^2 \tag{1.1.9}$$

Equation 1.1.9 gives the minimum norm estimator (MNE) as $\widehat{\mathbf{Z}} = (\mathbb{A}^T\mathbb{A} + \lambda\mathbf{I}_N)^{-1}\mathbb{A}^T\mathbf{X}$ or $\widehat{\mathbf{Z}} = \mathbb{A}^T(\mathbb{A}\mathbb{A}^T + \lambda\mathbf{I}_M)^{-1}\mathbf{X}$.

A more general MNE model sets the penalty term in equation 1.1.6 to be $f(\mathbf{Z}) = \lambda\|\mathbf{Z}\|_R^2$, which means there's no a priori information used, except that it assumes the current distribution has small overall intensity (L_2 penalty). The minimum norm estimator for the sources, denoted by $\widehat{\mathbf{Z}}$ can be computed as

$$\widehat{\mathbf{Z}} = \arg\min_{\mathbf{Z}} \|\mathbf{X} - \mathbb{A}\mathbf{Z}\|_Q^2 + \lambda\|\mathbf{Z}\|_R^2 \tag{1.1.10}$$

where $Q = \mathbb{C}\text{ov}\,(\mathbf{N}(t))$ and $R = \mathbb{C}\text{ov}\,(\mathbf{Z}(t))$.

The minimum norm estimator can be expressed explicitly as

$$\widehat{\mathbf{Z}} = R\mathbb{A}^T(\mathbb{A}R\mathbb{A}^T + \lambda Q)^{-1}\mathbf{X} \tag{1.1.11}$$

The minimum norm estimator (MNE) in favors weak and localized activation patterns. That means, superficial solution points in a 3D brain is favored since there is less activation needed to have similar impact on observations, comparing to some "deeper" sources.

- Weighted minimum norm (WMN) method

  In order to overcome the disadvantage of MN that it favors superficial sources, weighted minimum norm (WMN) (Gorodnitsky et al., 1995) approaches have been developed where different weighting strategies have been applied. The WMN estimator is com-

puted as

$$\widehat{\mathbf{Z}} = \arg\min_{\mathbf{Z}} \|\mathbf{X} - \mathbb{A}\mathbf{Z}\|_Q^2 + \lambda \|W\mathbf{Z}\|_F^2 \tag{1.1.12}$$

W can be determined in multiple ways. One possible choice for $W \in \mathbb{R}^N$ is to set W to be a diagonal matrix, where the diagonal elements are proportional to to the norm of the columns of the mixing matrix $\mathbb{A}$. That is

$$W_{nn} = \sqrt{\sum_{m=1}^{M} (\mathbb{A}_{mn})^2}.$$

- Laplacian weighted minimum norm (LORETA) method

  LORETA method sets the penalty term in Equation 1.1.6 to be $f(\mathbf{Z}) = \|DW\mathbf{Z}\|_F^2$, where $D \in \mathbb{R}^N$ is used to implement the discrete spatial Laplacian operator, W is a diagonal matrix for the column normalization of $\mathbb{A}$. The Laplacian matrix D can be written as

$$D_{ij} = \frac{1}{h^2} \begin{cases} -6 & \text{if} \|\mathbf{r}_i - \mathbf{r}_j\| = 0 \\ 1 & \text{if} \|\mathbf{r}_i - \mathbf{r}_j\| = h \\ 0 & \text{otherwise} \end{cases} \tag{1.1.13}$$

  where h is the minimum inter-grid distance. The LORETA estimator for $\mathbf{Z}$ can be obtained as

$$\widehat{\mathbf{Z}} = \arg\min_{\mathbf{Z}} \|X - A\mathbf{Z}\|_F^2 + \lambda \|DW\mathbf{Z}\|_F^2 \tag{1.1.14}$$

  LORETA in favors the solution with smooth spatial distribution, provided that it puts penalty on spatial roughness, which is measured by the Laplacian of the weighted sources. The assumption is related to the physiological in the sense that it assumes

neighboring sources are correlated. It also need to be pointed out that solutions obtained using LORETA are over-smoothed, in some situations.

- Local autoregressive average (LAURA) method

  LAURA approach (de Peralta Menendez et al., 2001) integrates the laws of physics in terms of a local spatial-temporal average with coefficients depending on the distance between solution points. LAURA incorporate biophysical laws by assuming that the electromagnetic activity follows Maxwell's law. The solution is obtained by

$$\widehat{\mathbf{Z}} = \arg\min_{\mathbf{Z}} \|\mathbf{X} - \mathbb{A}\mathbf{Z}\|_F^2 + \lambda \|WB\mathbf{Z}\|_F^2 \tag{1.1.15}$$

  where W is defined similarly as the diagonal matrix whose diagonal elements are the column norm of $\mathbb{A}$. The diagonal elements of B are defined as $B_{ii} = \frac{N}{N_i} \sum_{k \in V_i} d_{ki}^{-e_i}$, and the off-diagonal elements are zeros except for $k \in V_i$, where the value is defined as $B_{ik} = -d_{ki}^{-e_i}$, where $e_i = 2$, meaning that the fields are decreasing with the square of the inverse distance, d_{ik} is the Euclidean distance between point i and k, V_i represents the neighbor points of i, and N_i is the size of set V_i.

- Minimum current (MC) method

  Minimum current (MC) method applies L_1 norm penalty on source to obtain sparse solution (Matsuura and Okabe, 1995) (Uutela et al., 1999). In MC method, the penalty term in Equation 1.1.6 is defined as $f(\mathbf{Z}(t)) = \|\mathbf{Z}(t)\|_1$. For a given time t, the minimum current estimator for $\mathbf{Z}(t)$ can be obtained using

$$\widehat{\mathbf{Z}}(t) = \arg\min_{\mathbf{Z}}(t) = \|\mathbf{X}(t) - A\mathbf{Z}(t)\|^2 + \lambda \|\mathbf{Z}(t)\|_1 \tag{1.1.16}$$

The source reconstruction methods mentioned above are summarized in table 1.1.

Methods	Parameter and Signals	Dimensions	Objective function to be minimized	Remarks
Dipole fits	r_n, ϕ_n, $\mathbf{Z}_n$	$N < M$	$\|\mathbf{X} - \sum_{n=1}^{N} a(r_n, \phi_n)\mathbf{Z}_n\|_F^2$	estimate location and strength
LORETA	$\mathbf{Z}$	$N > M$	$\|\mathbf{X} - \mathbb{A}\mathbf{Z}\|_F^2 + \lambda\|DW\mathbf{Z}\|_F^2$	smooth solution
MN	$\mathbf{Z}$	$N > M$	$\|\mathbf{X} - \mathbb{A}\mathbf{Z}\|_Q^2 + \lambda\|\mathbf{Z}\|_R^2$	prefer surface/weak source
WMN	$\mathbf{Z}$	$N > M$	$\|\mathbf{X} - \mathbb{A}\mathbf{Z}\|_Q^2 + \lambda\|W\mathbf{Z}\|_F^2$	weighted according to depth
MC	$\mathbf{Z}(t)$	$N > M$	$\|\mathbf{X}(t) - \mathbb{A}\mathbf{Z}(t)\|_Q^2 + \lambda\|W\mathbf{Z}(t)\|_1$	sparse solution
LAURA	$\mathbf{Z}$	$N > M$	$\|\mathbf{X} - \mathbb{A}\mathbf{Z}\|_F^2 + \lambda\|WB\mathbf{Z}\|_F^2$	follow Maxwells' law

Table 1.1: Summary of Source Reconstruction Methods

1.1.4 Blind Source Separation

Generally, the source $\mathbf{Z}(t)$ and mixing matrix $\mathbb{A}$ are unknown in the model described in Section 1.1.2. Various methods have been developed to estimate the mixing matrix $\mathbb{A}$ and reconstruct the unknown source $\mathbf{Z}(t)$. Most commonly used methods in reconstructing the unknown sources are listed discussed below.

The inverse source reconstruction methods, as discussed in Section 1.1.3, are capable of estimating both the source activities and source locations. However, they rely on accurate forward individual head models and exact positions of electrodes. Inaccurate forward model may lead to inaccurate estimate. An alternative approach that does not depend on the head forward model estimates the mixing matrix $\mathbb{A}$ and source $\mathbf{Z}(t)$ in a completely data driven way. This class of approaches is referred to as blind source separation (BSS), which can be performed via a few approaches, including PCA, ICA and SOBI.

- Principal Component Analysis (PCA)

  According to (Jung et al., 2000), the mixing matrix $\mathbb{A}$ and source $\mathbf{Z}$ can be estimated through eigenvalue decomposition. Let $\mathbf{\Sigma} \in \mathbb{R}^{M \times M}$ be the covariance matrix of $\mathbf{X}(t)$, the PCA approach reconstructs the source by first performing eigen decomposition of

$\mathbf{\Sigma}$:

$$\mathbf{\Sigma} = WDW^T \tag{1.1.17}$$

where $W \in \mathbb{M} \times \mathbb{M}$ is unitary matrix, $D \in \mathbb{R}^{M \times M}$ is a diagonal matrix, with diagonal elements $d_1 \geq d_2 \geq$, ..., $\geq d_M$, which are eigen values of $\mathbf{\Sigma}$. The m-th column of W is the eigven vector of $\mathbf{\Sigma}$ corresponding to eigen value d_m. Let $W_N \in \mathbb{R}^{M \times N}$ be the matrix that is constructed by the first N columns of W, the reconstructed source $\widehat{\mathbf{Z}}(t)$ can be obtained using $\widehat{\mathbf{Z}}(t) = W_N^T \mathbf{X}(t)$. The mixing matrix $\mathbb{A}$ can be estimated using $\widehat{\mathbb{A}} = W_N$. The covariance matrix $\mathbf{\Sigma}$ in Equation 1.1.17 might be unknown, therefore we may use its estimate. An estimate of $\mathbf{\Sigma}$ can be computed as $\widehat{\mathbf{\Sigma}} = \frac{1}{T} \sum_{t=1}^{T} (\mathbf{X}(t) - \widehat{\mu})(\mathbf{X}(t) - \widehat{\mu})^T$, where $\widehat{\mu} = \frac{1}{T} \sum_{t=1}^{T} \mathbf{X}(t)$.

PCA looks for orthogonal directions along which the data has largest variance. The estimated sources are uncorrelated (second order independent) and orthogonal. It is reasonable to assume that the EEG sources are uncorrelated if the source are originally generated in different parts of the brain, however it is not reasonable in general to assume the EEG sources are spatially orthogonal. For this reason, PCA filtered signals are not generally interpreted as source signals. PCA is more commonly used in dimensionality reduction by projecting the signals onto a space with lower dimension. (Mørup et al., 2008) expanded the idea of SVD to higher order (tensor decomposition) so that it can handle multi-channel EEG data that is recorded for multiple trials.

- Independent Component Analysis (ICA)

  PCA de-correlate the data by diagonalizing empirical covariance matrix. For multivariate Gaussian, it is equivalent to removing the dependence between sources. However, for non-Gaussian signals, de-correlation does not necessarily imply removing dependency. (Makeig et al., 1996a) described an approach of using ICA to recover the source

signals in EEG records, and it is the first publication on ICA applied to EEG data. It recovers the source signal by minimizing the mutual information between source factors. ICA methods using mutual information depends on all higher-order statistics, comparing to PCA methods, which only depends on second order statistics. (Cardoso, 1999) describes one approach that recovers the independent components by Jacobi method of diagonalization. The measures of independence are based on fourth-order correlations between the observed signals.

According to (Makeig et al., 1996a), the ideal conditions for using ICA as a source separation technique include (1) sources are independent (2) propagation of sources to electrodes is negligible (3) observed EEG signals are linear mixing of sources. ICA also has assumption of non-Gaussianity of the source distributions. In practice, ICA should be restricted to data sets where the components show a significant amount of non-Gaussianity.

- Second Order Blind Identification (SOBI)

  An approach named Second Order Blind Identification (SOBI) has been developed in (Belouchrani et al., 1997). SOBI performs joint diagonalization of more than two cross-correlation matrices related to multiple time lags $\{\tau_1, \tau_2, ..., \tau_p\}$. The algorithm reconstruct $\mathbb{A}$ and $\mathbf{Z}$ by minimizing the cost function, which is the sum of the squares of off-diagonal elements of the transformed matrices (cross-covariance matrices for normalized signal). The assumption of SOBI requires that the source signals $\mathbf{Z}(t)$ is a stationary multivariate process that satisfies

  $$\mathbb{E}[\mathbf{Z}(t)\mathbf{Z}(t+\tau)^T] = diag[\rho_1(\tau), ..., \rho_N(\tau)]$$

  The advantage of using SOBI in EEG source reconstruction includes (1) it relies only on second order statistics (2) it allows separation of Gaussian sources and (3) it is more robust since it uses multiple cross-covariance matrices. The accuracy of SOBI

is affected by the overlapping of the spectra of sources. Large overlapping between spectra of sources may lead to inaccurate estimate.

1.2 Modeling and Inference Connectivity

The term connectivity may refer to structural, functional and effective connectivity (Friston, 1994). Where the structural connectivity refers to the anatomical structure of the brain, which can be studies via analysis of fMRI; the functional connectivity refers to undirected temporal correlations between neurophysiological events, which can be studies based on coherence or correlation; and the effective connectivity refers to the directed effects of one neural activity over another. Approaches to characterizing effective connectivity include Dynamic Causal Modeling (DCM)(Friston et al., 2003), Granger Causality Modeling (GCM) (Granger, 1988), and Transfer Entropy (Schreiber, 2000) and etc.

1.2.1 Coherence

- Coherence between two univariate time series.

  Coherence is the analog of cross-correlation in the frequency domain. Let $X(t) \in \mathbb{R}$ and $Y(t) \in \mathbb{R}$ be two time series which have zero mean and are jointly stationary. Then the cross-covariance between X and Y at lag h, denoted by $\gamma_{XY}(h)$, can be expressed as $\gamma_{XY}(h) = \mathbb{E}[X(t+h)Y(t)]$. The cross-spectrum $S_{XY}(\omega)$ is the Fourier transform of the cross-covariance γ_{XY} at frequency ω, which is in the form $S_{XY}(\omega) = \sum_{h=-\infty}^{\infty} \gamma_{XY}(h) \exp(-i2\pi\omega h)$, provided that $\sum_{h=-\infty}^{\infty} |\gamma_{XY}(h)| < \infty$. Then the coherence function is defined as

$$\rho^2_{XY}(\omega) = \frac{|S_{XY}(\omega)|^2}{S_{XX}(\omega)S_{YY}(\omega)}. \tag{1.2.1}$$

Coherence $\rho^2_{XY}(\omega)$ ranges from 0 to 1. It measures the strength of correlation and the phase consistency between X and Y at frequency ω.

- Coherence between multivariate time series: general coherence

  The concept of coherence has been generalized so that it can be computed for two multivariate time series. Let $X(t) \in \mathbb{R}^m$, $Y(t) \in \mathbb{R}^n$ denote two weakly stationary multivariate time series, where $X(t) = [X_1(t), \ldots, X_m(t)]'$, $Y(t) = [Y_1(t), \ldots, Y_n(t)]'$, $t = 1, 2, ..., T$. Let $S_{XX}(\omega) \in \mathbb{R}^{m \times m}$ be the spectral density matrix for $X(t)$ at frequency ω, where the (i, j)-th element of $S_{XX}(\omega)$ is defined as $S_{X_i X_j}(\omega)$. The spectral density matrix of $Y(t)$, denoted by $S_{YY}(\omega) \in \mathbb{R}^{n \times n}$, is defined in a similar way. The cross spectral matrix between $X(t)$ and $Y(t)$ at frequency ω is defined as $S_{XY}(\omega) \in \mathbb{R}^{m \times n}$, where the (i, j)-th element of $S_{XY}(\omega)$ equals to $S_{X_i Y_j}(\omega)$. According to (Pascual-Marqui, 2007b), the general coherence between two random vectors $X(t)$ and $Y(t)$ at frequency ω is defined as:

$$\rho^2_G(\omega) = 1 - \frac{\det[S_{YY|X}(\omega)]}{\det[S_{YY}(\omega)]} \tag{1.2.2}$$

  where $S_{YY|X}(\omega) = S_{YY}(\omega) - S_{YX}(\omega) S_{XX}(\omega)^{-1} S_{XY}(\omega)$.

  Remarks:

  - For univariate $X(t) \in \mathbb{R}$ and $Y(t) \in \mathbb{R}$, the general coherence given by Equation 1.2.2 will be equivalent to the coherence defined in Equation 1.2.1.
  - When $X(t)$ and $Y(t)$ are uncorrelated then $S_{YY|X}(\omega) = S_{YY}(\omega)$ and therefore $\rho^2_G(\omega) = 0$.

1.2.2 Partial Coherence

- Partial coherence between two univariate time series

  To obtain the direct (as opposed to indirect) dependence between two time series

$Y_1(t) \in \mathbb{R}$ and $Y_2(t) \in \mathbb{R}$, we need to remove the linear effect of all other time series from each of $Y_1(t)$ and $Y_2(t)$. One way to measure this is via partial coherence.

To illustrate this concept, consider the multivariate time series $Y(t) = [Y_1(t), Y_2(t), Y_3(t)]$. Define the spectral matrix to be

$$\boldsymbol{S}(\omega) = \begin{pmatrix} S_{Y_1Y_1}(\omega) & S_{Y_1Y_2}(\omega) & S_{Y_1Y_3}(\omega) \\ S_{Y_2Y_1}(\omega) & S_{Y_2Y_2}(\omega) & S_{Y_2Y_3}(\omega) \\ S_{Y_3Y_1}(\omega) & S_{Y_3Y_2}(\omega) & S_{Y_3Y_3}(\omega) \end{pmatrix}. \tag{1.2.3}$$

Define the matrix $\Gamma(\omega)$ to be $\Gamma(\omega) = -\boldsymbol{M}(\omega)\boldsymbol{H}(\omega)\boldsymbol{M}(\omega)$, where $\boldsymbol{H}(\omega) = \boldsymbol{S}^{-1}(\omega)$ and $\boldsymbol{M}(\omega)$ is a diagonal matrix whose elements are $\frac{1}{\sqrt{H_{nn}(\omega)}}$. The partial coherency between time series Y_1 and Y_2 at frequency ω is the $(1,2)$ element of the matrix $\Gamma(\omega)$ denoted $\Gamma_{12}(\omega)$. Partial coherence between Y_1 and Y_2 at frequency ω is $\lambda_{12}(\omega) = |\Gamma_{12}(\omega)|^2$. Partial coherence $\lambda_{12}(\omega)$ ranges from 0 to 1. It measures the coherence between signals X and Y after removing linear effect of Z.

- Partial coherence between multivariate time series: Block partial coherence

  Consider multi-variate time series $X(t)$, $Y(t)$ and $Z(t)$, according to (Nedungadi et al., 2011), the block partial coherence at frequency ω between X and Y, conditional on Z is defined as

$$C^{(B)}_{XY|Z}(\omega) = 1 - \frac{\det[S_{[X,Y]|Z}(\omega)]}{\det[S_{XX|Z}(\omega)]\det[S_{YY|Z}(\omega)]}$$

  If $X(t)$, $Y(t)$ and $Z(t)$ are all uni-variate time series, $C^{(B)}_{XY|Z}(\omega)$ is reduced to the partial coherence for the uni-variate time series.

1.2.3 Granger Causality

The concept of Granger causality (GC) was first proposed in (Granger, 1969a). The basic idea of GC states that for random variable $X(t)$ and $Y(t)$ with $t = \{1, \ldots, T\}$, we say X "Granger-causes" (GC) Y if Y can be better predicted using the past value of both X and Y than it can using the past observation of Y alone. According to (Granger, 1980) the causality relationship are based on two principles: (1.) the cause must happen prior to to its effect (2.) the cause must contain unique information for predicting future values of the effect.

The limitation of Granger causality includes: (1) It does not necessarily mean physiological causality; (2) GC can only account for linear relationship; (3) GC relies on the stationarity of the analyzed the signals. Nevertheless, we shall explore its utility in this study given that it is widely used in neuroscience.

To illustrate the idea of Granger-causality, consider the following two auto-regressive models

$$X(t) = \sum_{k=1}^{p} \Phi_k X(t-k) + \epsilon(t) \tag{1.2.4}$$

$$X(t) = \sum_{k=1}^{p} \Phi'_k X(t-k) + \sum_{k=1}^{p} \Psi'_k Y(t-k) + \epsilon'(t) \tag{1.2.5}$$

If model 1.2.5 is significantly better than 1.2.4, that is

$$\mathbb{E}\Big[\sum_{t=P+1}^{T} (X(t) - \sum_{k=1}^{p} \Phi_k X(t-k))^2\Big] > \mathbb{E}\Big[\sum_{t=P+1}^{T} (X(t) - \sum_{k=1}^{p} \Phi'_k X(t-k) - \sum_{k=1}^{p} \Psi'_k Y(t-k))^2\Big]$$

then we can conclude that Y Granger-causes (GC) X.

1.2.4 Partial Directed Coherence (PDC)

Partial directed coherence (PDC) has been introduced in making inference on frequency specific connectivity between signals(Baccalá and Sameshima, 2001a). It is a directed measurement of connectivity and is based on the vector autoregressive model of the signals. Partial directed coherence can be treated as frequency description of Granger-causality.

Let $X(t) = [X_1(t), \ldots, X_M(t)]$ be a stationary M dimensional time series with mean zero. A vector autoregressive (VAR) model with order p for $X(t)$ is given by

$$X(t) = \sum_{k=1}^{p} \Phi_k X(t-k) + E(t) \tag{1.2.6}$$

where $\Phi_k \in \mathbb{R}^{M \times M}$ is the coefficient matrix of the VAR model at lag k; $E(t) \in \mathbb{R}^M$ is a Gaussian white noise process with with zero mean and covariance matrix Σ_E. The causality of $X(t)$ requires the coefficient matrices satisfy $\det(I - \sum_{k=1}^{p} \Phi_k z^k) \neq 0$ for all $z \in \mathbb{C}$ and $|z| \leq 1$.

Let $\Phi(\omega) = I - \sum_{k=1}^{p} \Phi_k \exp(-i2\pi\omega k/\Omega_s)$ be the Fourier transform of Φ_k at frequency ω, where Ω_s represents the sampling frequency. Then the partial directed coherence $\pi_{ij}^2(\omega)$ for the VAR process defined in Equation 1.2.6 can be represented as

$$\pi_{ij}^2(\omega) = \frac{|\Phi_{ij}(\omega)|^2}{\sum_{k=1}^{N} \Phi_{kj}(\omega)\Phi_{kj}^*(\omega)} \tag{1.2.7}$$

$\pi_{ij}^2(\omega)$ provides a measure of the linear influence of X_j on X_i at frequency ω. The partial directed coherence $\pi_{ij}^2(\omega)$ takes values from interval $[0, 1]$ and satisfies $\sum_i \pi_{ij}^2(\omega) = 1$, which means each time series has unit outflow.

The connectivity measures mentioned above are summarized in table 1.2.

Approach	Model Based	Data Driven	Functional Connectivity	Effective Conenctivity
DCM	✓			✓
Fourier Coherence		✓	✓	
Wavelet Coherence		✓	✓	
Granger Causality		✓	✓	✓
PDC		✓	✓	✓
DTF		✓	✓	✓

Table 1.2: Summary of Connectivity Measures

1.3 Outlines

The remainder of this dissertation is organized as follows. Chapter 2 explores the approaches for analyzing high-dimensional time series through low dimensional embeddings. Chapter 3 introduces the proposed source-space factor VAR model and its application on making inference on cortical connectivity from EEG observations. The last chapter, Chapter 4 discusses the conclusions and future directions.

Chapter 2

Time Series Low Dimensional Embeddings

In this chapter, we develop computationally efficient and theoretically justified tools for analyzing high dimensional brain signals. Our approach is to extract the optimal lower dimensional representations for each brain region and then characterize and estimate connectivity between regions through these factors. This approach is motivated by our observation that electroencephalograms (EEGs) from many channels within each region exhibit a high degree of multicollinearity and synchrony thereby suggesting that it would be sensible to extract summary factors for each region. Here, the summary factors are actual solutions to squared error reconstruction error. We focus on two special cases of linear auto encoder and decoder. The first characterizes the factors as instantaneous linear mixing of the observed signals. In the second approach, the factors are convolutions of the observed signals (which is more general than the first). We implemented these methods in a Matlab toolbox XHiDiTS (`https://goo.gl/uXc8ei`). These methods were compared through simulations under different conditions and the results provide insights on advantages and limitations of each. Finally, we performed exploratory analysis of resting state EEG data and fMRI data. The

spectral properties of the factors were estimated and connectivity between regions via the factors using coherence measures were computed. The toolbox was utilized to investigate consistency of these factors across all epochs during the entire resting-state period.

2.1 Introduction

Neuronal populations behave in a coordinated manner both during resting-state and while executing tasks such as learning and memory retention. One of the major challenges to analyzing brain signals such as electroencephalograms (EEGs) and functional magnetic resonance imaging (fMRI) is high dimensionality. For example, there can be hundreds of channels in a typical EEG recording, and the number of voxels in a fMRI recording can be hundreds of thousands. Motivated by the typical fMRI analysis, our approach is to parcellate the cortex into small set of regions based on some prior information of functional and anatomical proximity, and then analyze clusters of signals within and between these regions. Depending on the parcellation of the brain cortex, a cortical region could correspond to 15-25 EEG channels (Wu et al., 2014). Empirical inspection of the resting-state EEGs in each region in Figure 2.1 shows a high degree of collinearity and synchronicity. EEGs, which indirectly measure electrical activity in the cortex, show high multicollinearity between channels due to spatial filtering and volume conduction. Thus, it is sensible to reduce dimensionality in each region by deriving some signal summaries. A naïve solution is to compute the average across all channels. As we will demonstrate in this chapter, simple averaging is problematic especially when some of the EEGs are out of phase. Here, we will use a data-adaptive filtering (encoding) procedure to extract representative signals that is a solution to minimizing the squared reconstruction error of the original high-dimensional signals.

Obtaining summaries for each brain region is the first step towards the next goal which is to study connectivity between brain regions. Connectivity has been a focus of many studies both

in cognitive science and clinical neuroimaging. In fact, disruptions in connectivity between brain regions is associated with a number of neurological diseases such as schizophrenia, obsessive compulsive disorder and Alzheimer's disease. The common measure for brain functional connectivity is correlations between multivariate signals measured from the brain regions. The challenge when analyzing connectivity between large number of regions involves estimation of a high-dimensional correlation structure. The signal dimensionality is often comparable to or even larger than the sample size, where traditional estimators are known to produce unreliable estimates.

Our approach to modeling and estimating the connectivity between brain regions is to first derive low dimensional representation or simpler structure of the original signals based on factor analysis for each brain region. We then study connectivity structure between brain regions via the derived lower dimensional factor space. The proposed general framework is illustrated in Figure 2.1. The focus of this chapter is on the first step of the approach and to conduct an exploratory study of two general procedures for extracting a small number of latent factors in multichannel EEGs. The first assumes the factors as an instantaneous mixing and the second a linear filtering of the high-dimensional signals. This is a first step towards dimension reduction which could lead to better statistical modeling and inference. It could help to identify any potential irregularities in the signals (e.g. outliers, non-stationarities across epochs). These derived summaries are useful for capturing the hidden features in high dimensional signals. Moreover, the factors and the mixing matrix derived in the exploratory analyses can be use to simulate realistic data that closely resemble real EEG signals. Finally, since these EEG signals are recorded over several epochs, one should not ignore the possibility that the underlying brain process has evolved across the entire experiment. Thus, one of our goals is to develop a tool that explores how connectivity patterns might have evolved during the experiment.

As noted above, the most natural approach to summarizing activity in one brain region

is to simply take the average signal. In fact, connectivity analyses of functional magnetic resonance imaging (fMRI) are usually conducted by taking the fMRI time series average across voxels in a pre-defined region of interest. (See (Fiecas et al., 2013; Gott et al., 2015) and references cited). While this approach is certainly intuitive and easy to implement, it has a number of serious limitations. In fact, (Sato et al., 2010) already pointed the pitfalls and suggests a data-driven approach via conventional principal components analysis (PCA) which is essentially an instantaneous (or contemporaneous) mixing of time series. Moreover, (Lila et al., 2016) developed a regularized PCA approach for functional data on a two-dimensional manifold, where the regularization is applied on the curvature of the function over the manifold. Also, (Lazar, 2008) discussed general concepts of handling high dimensionality in fMRI via principal components. In order to deal with the high dimensionality of fMRI data and complexity of the spatio-temporal covariance structure, (Kang et al., 2012) proposed a procedure that transforms the (residual) fMRI time series in each voxel into the Fourier domain. The rationale is that, under weak stationarity, the Fourier coefficients are approximately uncorrelated (Shumway and Stoffer, 2010). Other approaches for modeling brain connectivity from high-dimensional brain imaging data include Dynamic Connectivity Regression (DCR) (Cribben et al., 2012), Dynamic Conditional Correlation (DCC) (Lindquist et al., 2014), group independent component analysis (ICA) (Calhoun et al., 2014; Calhoun and Adali, 2012), and sparse vector autoregressive (VAR) modeling (Davis et al., 2015).

In this work, we will use a more general approach where the factor-based signal summaries are produced by filtering and hence utilizes temporal correlation structure of the data at different lags rather than just the zero-lag. The factors are extracted by using spectral PCA. In contrast to (Lila et al., 2016) we focus on capturing the spectral property of signals like EEGs which are often not smooth and are characterized by oscillatory and high frequency processes. This poses greater challenges than analyzing fMRI time series which are dominated by low-frequency components with short dependence (often assumed to follow

first-order autoregression). Our approach offers two advantages over other representative signals in addressing these problems. First, due to its complexity, the underlying features of a brain process is unlikely to be captured by a single summary such as the average. Second, due to potential phase-shifts and lead-lag dependence among the channels, the averaged signal could produce misleading results: averaging perfectly out-of-phase signals produces "flat lines" and averaging shifted signals could produce higher-frequency artifacts. This is discussed further in Section 2.3.2. On the other hand, factors extracted by conventional PCA can account for the signal variability better than the simple averages. However, it is less effective in capturing the lagged dependence compared to the spectral PCA.

The remainder of the chapter is organized as follows. Section 2.2 describes the methods for dimensionality reduction for time series data, connectivity measurements and our implementation of the methods as an interactive Matlab toolbox. Section 2.3 presents the evaluation results of the models based on simulated data and the results of an exploratory analysis on a real resting-state EEG dataset and a fMRI dataset. In Section 2.4, we summarize the conclusions and propose future directions.

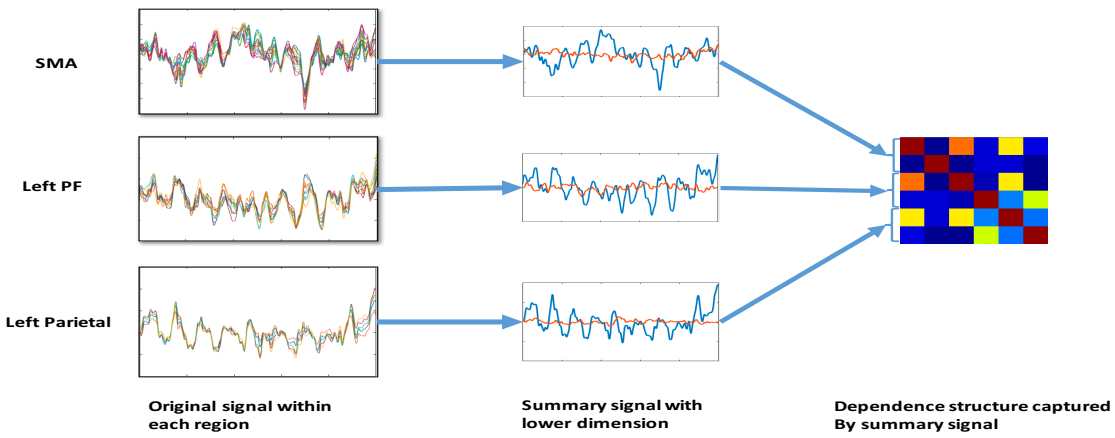


Figure 2.1: Illustration of the exploratory procedure. On the left are the EEG signals from three different regions: supplementary motor area (SMA), left Pre-frontal cortex and left parietal. The goals are (1.) to obtain summaries within each region and (2.) to compute dependence between the brain regions through the factors.

2.2 Methods

In this section, we describe our approach to performing exploratory analysis of high-dimensional EEG time series. Let R be the number of regions on the scalp topography. Denote the EEG signals in a region $r \in \{1, \ldots, R\}$ to be $\mathbf{z}_r(t)$ which we assume to have dimension n_r (i.e., the number of channels within region r). Recall that EEG signals within a region appear to be highly correlated which indicates that the EEGs can be represented by a more compressed time series (i.e., the factors) with lower dimension. There are advantages to low dimensional representation or embedding. (1.) The simplified structure helps in interpretation of the results. (2.) It takes advantage of the structure of data (e.g., high correlation for EEG signals within the same region) and hence the low dimensional embedding serves a similar purpose as imposing regularization. (3.) It enables the modeling estimation using a relatively small training sample size-reduced number of parameters enables us to fit a less constrained parametric model with better accuracy, for example, allowing larger lags in vector autoregressive (VAR) model allows us to capture more complex dependence structure; and (4.) Finally, models with reduced dimension can be trained more efficiently requiring fewer signals in the training dataset. This can significantly reduce the computational time especially when the training time is higher than quadratic order of the dimension. For example, in computing partial coherence, the inversion operation is faster and will have fewer problems resulting from matrix inversion since the dimension is lower.

Substantial research has been established on dimensionality reduction of high dimensional time series. Such methods include time-domain and frequency-domain PCA (Brillinger, 1964, 2001; Wang et al., 2016b; Stock and Watson, 2002); factor models (Durbin and Koopman, 2001; Harvey, 1990; Lam and Yao, 2012; Bai and Ng, 2002; Forni et al., 2012); canonical analysis (Box and Tiao, 1977), and independent component analysis (Samworth et al., 2012). . Here, we build on these foundations in two directions: we develop exploratory tools that handle factor analysis of partitioned signals (rather than all signals as a whole) with

extension to multiple-epochs (rather than just a single epoch). We then package these into a toolbox that we hope would help neuroscientists to use a more data-adaptive approach to investigating connectivity in high dimensional brain signals. The methods are based on learning compressed representations of the original time series of higher dimension. There are similar toolboxes which however focused on regularized estimation of high-dimensional time series in the original space (e.g., the big-VAR model (Nicholson et al., 2014) and the sparse-VAR model(Basu et al., 2015)).

This chapter will be mainly focused on two classes of approaches for extracting low dimensional summaries both of which are based on principal component analysis or factor analysis in a more general sense. We now describe the algorithm for deriving the factor activities for each region r, denoted as $\mathbf{f}_r(t)$, which has dimension m_r that is much smaller than that of the original signal n_r (i.e., $m_r \ll n_r$). For consistency in notation, we use a lower case letter for a scalar number; bold lowercase letter for a column vector; and uppercase letter for a matrix. When it is clear from the context, we shall drop the subscript r when dealing with the time series signals.

2.2.1 Auto encoder for time series data

The auto encoder algorithm is a general approach to learning compressed representations of the input data (e.g., high dimensional time series). The general algorithm was first introduced by (Rumelhart et al., 1985) and can be used to reduce the dimension of the time series via learning a low dimensional representation. The algorithm consists of two parts, namely, the encoder and the decoder. The *encoder* function $F_{\text{en}} : \mathcal{Z} \to \mathcal{F}$ is a mapping from the original high-dimensional space $\mathcal{Z}$ to space $\mathcal{F}$ of lower dimension. The *decoder* function, defined as $F_{\text{de}} : \mathcal{F} \to \mathcal{Z}$ is a mapping from the encoded (low-dimensional) space $\mathcal{F}$ to the original high dimensional space $\mathcal{Z}$.

Consider time series $\mathbf{z}(1), \ldots, \mathbf{z}(T)$ generated from process $P(Z)$, the optimal mapping F_{de} and F_{en} are the ones that minimize the expected reconstruction error defined as

$$L(F_{\text{de}}, F_{\text{en}}) = \mathbb{E}_{P(Z)} ||Z - F_{\text{de}}(F_{\text{en}}(Z))||_F^2 \tag{2.2.1}$$

where $|| * ||_F$ is the Frobenius norm, which is defined as $||E||_F = \sqrt{\text{Trace}(EE')}$.

We will consider special cases where the encoders and decoders are linear transformations (either instantaneous mixing or filtered versions) of the original time series. For linear encoder and decoders, i.e., both of F_{en} and F_{de} are linear functions, the solution is closely related to principal component analysis under the Frobenius norm based on the squared error of reconstruction. In this work, we will focus on two types of linear encoders and decoders: the first is the instantaneous mixing encoder and the second is a linear filter of the time series. The latter captures the entire temporal dynamics (lead-lag structure) of the time series.

2.2.2 Method 1: The factor is an instantaneous linear mixture of the time series $\mathbf{z}(t)$

For a zero-mean stationary time series process $\mathbf{z}(t) \in \mathbb{R}^n$ (we drop the subscript r for ease in presentation), we consider the problem of learning a representation of lower dimension. For the encoding function, we consider the mapping where the factor $\mathbf{f}(t) \in \mathbb{R}^m$ is represented as the instantaneous linear mixture of the original time series $\mathbf{z}(t)$. That is, $\mathbf{f}(t) = A'\mathbf{z}(t)$. Similarly, for the decoder function, we consider the mapping $\mathbf{z}(t)$ is the instantaneous linear mixture of $\mathbf{f}(t)$ in the form $\widehat{\mathbf{z}}(t) = B\mathbf{f}(t)$. For the purposes of keeping the parameters identifiable, we shall assume that (a.) $A'A = I_m$ and (b.) $\mathbb{C}\text{ov}[\mathbf{f}(t)]$ is a diagonal matrix, i.e., the factors are uncorrelated. The optimal representation is considered as the one that

gives the best reconstruction accuracy, which can be derived using the following two steps.

- Step 1. Compute the eigenvalues-eigenvectors of $\Sigma^{\mathbf{z}}(0)$ as $\{(\lambda_s, \mathbf{e}_s)\}_{s=1}^n$ where $\lambda_1 > \ldots, > \lambda_n$ and $\|\mathbf{e}_s\| = 1$. When $\Sigma^{\mathbf{z}}(0)$ is not known, we use an estimator instead, which can be computed as $\widehat{\Sigma}^{\mathbf{z}}(0) = \frac{1}{T}\sum_{t=1}^T \mathbf{z}(t)\mathbf{z}(t)'$ assuming $\mathbf{z}(t)$ has zero mean.

- Step 2. The solution can be represented by

$$\widehat{A} = \widehat{B} = [\mathbf{e}_1, \ldots, \mathbf{e}_m] \quad \text{and} \quad \widehat{\mathbf{f}}(t) = \widehat{A}'\mathbf{z}(t) \ .$$

The solution is identical to principal components analysis (PCA) on the input signals , using the covariance matrix of lag zero, i.e., $\Sigma^{\mathbf{z}}(0) = \mathbb{C}\text{ov}(\mathbf{z}(t), \mathbf{z}(t))$. It is the one that accounts for the most of the variation of the time series, among all the instantaneous linear projections with the same dimension.

2.2.3 Method 2: The factor is a linear filter of the time series $\mathbf{z}(t)$

Alternative to instantaneous mixing method, we consider a more flexible linear encoding and decoding functions where the encoding function is a linear convolution of all $\mathbf{z}(t) \in \mathbb{R}^n$ – rather than merely an instantaneous linear mixture. The lower dimensional representation (encoder function) can be written as

$$\mathbf{f}(t) = \sum_{h=\infty}^{\infty} A(h)'\mathbf{z}(t-h) \tag{2.2.2}$$

where $A(h) \in \mathbb{C}^{n\times m}$ with $m < n$, and $f_i(t)$ and $f_j(t)$ has zero coherency (in other words, uncorrelated at all lags) for $i \neq j$. For the reconstruction (decoder) function, we consider

the following form

$$\widehat{\mathbf{z}}(t) = \sum_{j=-\infty}^{\infty} B(j)\mathbf{f}(t-j) \tag{2.2.3}$$

where $B(j) \in \mathbb{C}^{n\times m}$ is the transformation coefficient matrix. The optimal values of $A(h)$ and $B(j)$ are chosen to minimize the reconstruction error defined in Equation 2.2.1. The solution can be obtained via principal components analysis of the spectral matrix. Let $S_{zz}(\omega)$ be the cross spectral density for $\mathbf{z}(t)$ at frequency ω. Denote the eigenvalue of $S_{zz}(\omega)$ as $\{\lambda_1 > \lambda_2, \ldots, \lambda_n\}$; denote the corresponding eigenvectors as $\{\mathbf{e}_1(\omega), \mathbf{e}_2(\omega), \ldots, \mathbf{e}_n(\omega)\}$, then the solution can be obtained as

$$A(h) = \int_{-1/2}^{1/2} A(\omega)\exp(i2\pi h\omega)d\omega \tag{2.2.4}$$

$$B(j) = \int_{-1/2}^{1/2} B(\omega)\exp(i2\pi j\omega)d\omega \tag{2.2.5}$$

where $B(\omega) = \overline{A(\omega)}$ and $A(\omega) = [\mathbf{e}_1(\omega), \ldots, \mathbf{e}_m(\omega)]$.

The detailed algorithm is described in Algorithms (1) and (2). Note that this dimension reduction procedure was originally described in (Brillinger, 1964). In this chapter, we shall refer to this as the "Spectral-PCA" method.

2.2.4 Comparison of the encoding/decoding methods

Both encoder methods are based on projecting the original high dimensional signal onto a space of lower dimension. Both methods are similar in the sense that the factors are constrained to be linear functions of the original signal. However they differ in this respect: the instantaneous mixing method produces factor $\mathbf{f}(t)$ which only explicitly depends on the

signal at time t. Under this approach, temporal dynamics of $\mathbf{z}(t)$ is ignored. The spectral-PCA method gives factors which are low dimensional filtered versions of the original signal. The factors at time point t is obtained by using *all* data points $\mathbf{z}(t \pm \ell)$. Thus it captures temporal dynamics and lead-lag relationships in the original time series. We note here that the first method is a special case of the second. In fact, by constraining $A(h) = 0$ and $B(j) = 0$ for all $h \neq 0$ and $j \neq 0$ then the linear filtered series is reduced to the instantaneously-mixed signal.

The key advantage of the spectral-PCA method is that it is likely to give lower reconstruction error because it uses all the information about the signal. The first method is particularly problematic when there is some lead-lag relationships between the original signals which could be completely washed out with the simplistic instantaneous mixing. It is also supported by the simulation results, where the spectral-PCA method has better performance when the time series has time shift in some channels (Figure 2.6 and 2.7).

In terms of computational complexity, the spectral-PCA needs to compute the eigenvalue decomposition of the spectral matrices *all* the frequencies while the instantaneous mixing method only needs to decompose the zero-lag covariance matrix. That means the second method requires more computational resources (in both space and time), comparing to method 1. It would be helpful to identify the method with suitable complexity for the problem. In the simulation study, we applied the two models on time series data generated from different distributions to gain better understanding of their performance.

A similar approach named generalized dynamic principal component analysis (GDPCA) has been proposed in (Peña and Yohai, 2015). Unlike method 1 and method 2 where information of zero lag or infinite lags are used, the encoding function for the GDPCA is a linear filter of factors with finite lags, and thus the method can be adapted to non-stationary time series as well. In GDPCA, there is no constraint of the form of the encoding function, therefore the encoding can be non-linear. Unlike method 1 and method 2 where the explicit solution

can be computed, in GDPCA, the solution is computed through an iterative algorithm and there is no explicit form of the encoding function.

2.2.5 Measures of Connectivity

The brain connectivity may refer to structural, functional and effective connectivity (Friston, 1994), where the structural connectivity refers to the anatomical structure of the brain; functional connectivity refers to undirected temporal correlations between neurophysiological events and effective connectivity refers to the directed effects of one neural activity over another. The effective connectivity measure by dynamic causal modeling (DCM)(Friston et al., 2003), Granger causality modeling (GCM) (Granger, 1988), transfer entropy (Schreiber, 2000), partial directed coherence (PDC) and isolated effective coherence (iCoh). Functional connectivity can be measured using coherence, generalized coherence and partial coherence. Due to the identifiability issue of the sign of the factors, we use the magnitude of the connectivity measures. This work will be focused on the functional connectivity between multivariate time series, specifically, the coherence between two time series. We demonstrate the connectivity estimation by applying the methods on resting-state EEGs recorded in a motor learning study. EEG signals from three regions (i.e., supplementary motor area (SMA), left lateral parietal region (left_latPr) and left medial parietal region (left_medPr)) were selected and the factors for each region were obtained via spectral PCA. Then we characterized the connectivity by computing the coherency between EEG channels and between the factors. The results displayed in Figure 2.2 are as expected, where the coherence between left_latPr and left_medPr at 10 Hz is stronger than other region pairs, which can be shown in the coherence estimate for the EEGs and can be summarized based on the factors.

2.2.6 Interactive Matlab toolbox for exploratory analysis

We implemented and actively maintain a Matlab toolbox (Exploratory High-Dimensional Time Series (XHiDiTS) toolbox `https://goo.gl/uXc8ei`) with a graphical interface that allows users to performance exploratory analysis easily. Figure 2.2 shows a screen shot of the toolbox interface.

The option panels provide a rich set of options that allow users to select from by just one-clicking. The options include (1.) subject-specific data; (2.) experimental conditions (resting state vs task); (3.) specific regions (users can load their own channel location/ grouping files) (4.) methods for learning lower dimensional representations and the complexity of the model (e.g. number of factors); and (5.) methods for computing connectivity (e.g. partial directed coherence, correlation matrix, coherence matrix and block coherence).

The visualization panels show (1.) the 2-d scalp, with selected regions highlighted and colored, where the coloring is consistent with the title of the signal plot, allowing users to match the plot and region easily; (2.) the signals and factors (low dimensional representations) for selected regions; and (3.) the spectrum of the signals and the connectivity maps.

The toolbox has low latency in updating the results for datasets with reasonable sizes. For example, for a 256-channel EEG data that contains 1000 time points for each channel and 200 epochs, the latency for updating the results for a new setup is within seconds (¡1s for most of the methods). Users can also load their own datasets or add their own definition of functions for connectivity and other quantities.

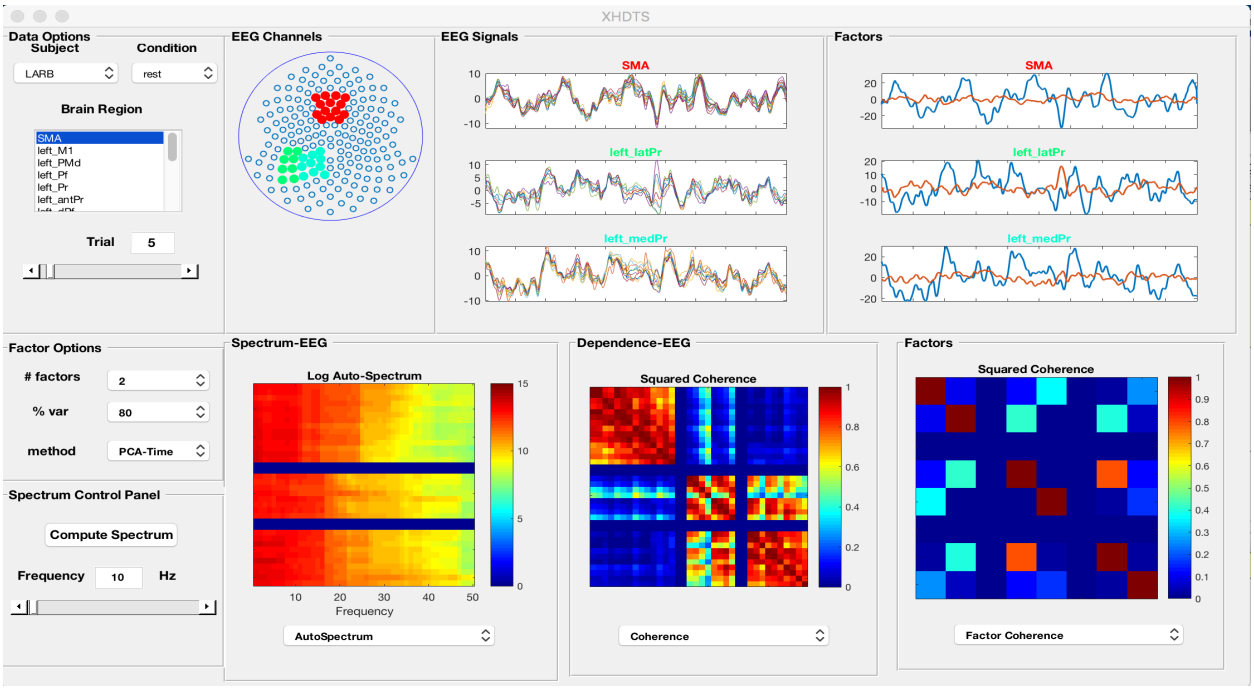


Figure 2.2: The interface of the XHiDiTS toolbox: exploratory high dimensional time series toolbox in Matlab. This is an interactive toolbox where the user selects the dataset to be analyzed. From the dataset, the user can select specific regions of interest (ROIs) to be analyzed. This toolbox supports a rich set of options and methods for visualizing and analyzing high dimensional time series, including the methods presented in this chapter. This toolbox is actively developed and maintained. It can be downloaded from `https://goo.gl/uXc8ei`. In this view, the top row includes (a.) channels on selected regions (SMA, left_latPr and left_medPr), (b.) EEG time series on the 3 different regions and (c.) summary factors time series of each region; bottom row includes (a.) the EEG power spectrum density, (b.) the coherency between EEG channels and (c.) the coherency between the corresponding factors.

2.3 Results and Discussion

2.3.1 Simulation

In this section, we apply the methods on simulated time series of various properties and evaluate the performance of the proposed approaches for obtaining signal summaries. The goal of these simulation studies is to provide a comprehensive evaluation of the methods

including computational complexity, accuracy (the squared reconstruction error criterion) and the application scope. The steps of the simulations are described as follows.

- Step 1: model training. Generate training time series data Z_{tr} from ground-truth generating process $P(Z)$, and then fit the model using methods 1 and 2 based on Z_{tr}.
- Step 2. model evaluation. Generate K iid test datasets $Z_{te_1}, \ldots, Z_{te_K}$ from $P(Z)$. For each test dataset Z_{te}, we evaluate the model using the normalized reconstruction error in the form of $\frac{||Z_{te}-\widehat{Z}_{te}||_F^2}{||Z_{te}||_F^2}$, where the reconstructed time series $\widehat{Z}_{te}$ is computed by applying the trained model obtained in Step 1. The mean and the standard deviation of the test error are computed over the K test sets and compared across models.

We performed multiple simulations using data generated from different ground-truth data generating processes of different complexities. We first generate the source signal $\mathbf{f}(t) \in \mathbb{R}^m$ for $t = 1, \ldots, T$. The observations are generated via

$$\mathbf{z}(t) = \sum_{h=0}^{t-1} Q(h)\mathbf{f}(t-h) + \mathbf{e}(t) \tag{2.3.1}$$

We will demonstrate here the advantages and potential pitfalls of the methods under different application settings, by varying $\{Q(h)\}$ and the distribution of $\{\mathbf{f}(t)\}$. The comparisons between two methods are performed on the time series of different spatial and temporal correlation structures. The simulation setups are summarized in Table 2.1.

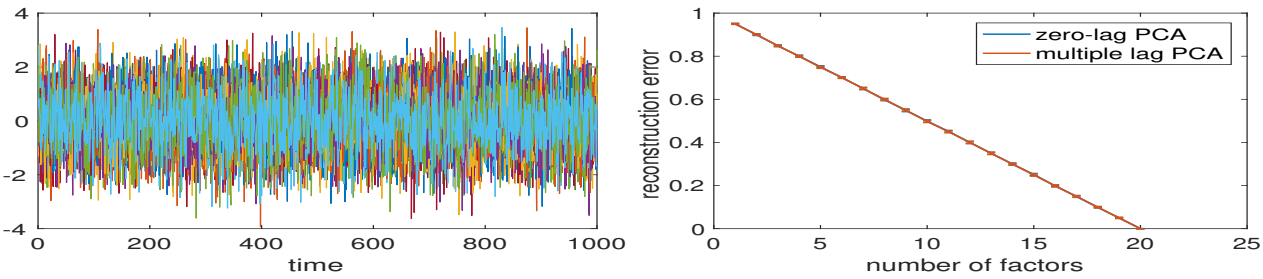
Setting	$\mathbf{f}(t)$	$Q(h)$	spatial corr	temporal corr	phase shift
1.	iid $N(0, I_{20})$	$Q(0) = I_{20}$	no	no	no
2.	iid $N(0, I_2)$	$Q(0) \in \mathbb{R}^{20\times 2}$	yes	no	no
3.	AR(1)	$Q(0) \in \mathbb{R}^{20\times 1}$	yes	yes	no
4.	AR(2)	$Q(0)[11:20,1] \in \mathbb{R}^{10\times 1}$ $Q(40)[1:10,1] \in \mathbb{R}^{10\times 1}$	yes	yes	yes
5.	AR(2)	$Q(80)[7:12,1] \in \mathbb{R}^{6\times 1}$ $Q(40)[1:6,1] \in \mathbb{R}^{6\times 1}$ $Q(0)[13:20,1] \in \mathbb{R}^{8\times 1}$	yes	yes	yes

Table 2.1: Summary of the simulation set up. $Q(h)[i:j,k]$ denotes the elements located in the k-th column and between row i and j (inclusive) of matrix $Q(h)$. For the value of $Q(h)$, unless it is explicitly defined in the table, is set to 0 by default.

Setting 1: the multivariate time series are spatially and temporally independent

In this setting, we consider the distribution of the data $P(Z)$ to be spatial independent and temporal independent. In this case, the signals are spatially independent and therefore the dimension reduction is less effective. Moreover, since the factors are also all temporally independent, these summaries should all reflect white noise properties. Our simulation studies indeed verify these expected outcomes. It can be observed form Figure 2.3 that comparing between two methods, the reconstruction errors evaluated on the test datasets are similar. For both models, the reconstruct error appears to decrease linearly as the number of factors increases, which indicates that in the iid Gaussian case, all factors account for the same amount of the total variation. The result is as expected because (1.) for iid Gaussian random variable with identity covariance matrix, the projection on any direction will account for the same amount of variation and (2.) there is no lead-lag relationship between observations at different time points therefore in terms of predicting current observation, there is no gain of

using observations from past or in the future. It can also be shown that in theory, method 1 and 2 will have the same solution when the signals are iid Gaussian with zero mean and identity covariance matrix.



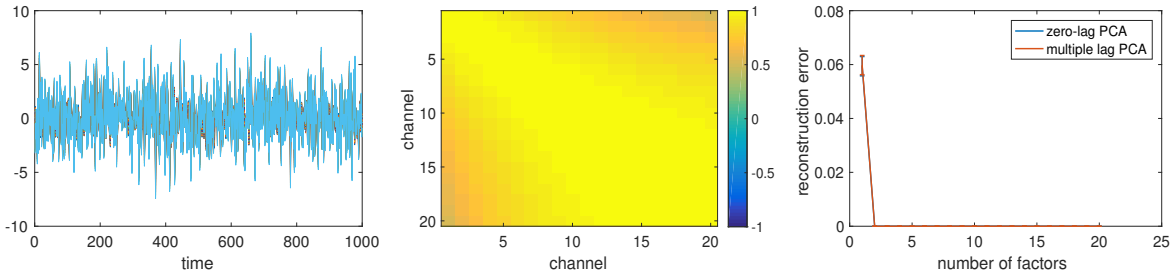
(a) Plot of $\mathbf{z}_{tr}(t)$

(b) Reconstruction error

Figure 2.3: Independent time series. The reconstruction error (squared loss, i.e., the square of the Frobenius norm of the residuals) appears to be decreasing linearly when number of factors increases which means in the iid Gaussian case, all factors account for the same amount of the total variation.

Setting 2: the multivariate time series are temporally independent but spatially correlated

In this setting, we consider white noise time series where $\mathbf{z}(t)$ and $\mathbf{z}(t+h)$ are independent Gaussian variables for $h \neq 0$. With this setting, the covariance $\mathbb{C}\text{ov}\,\mathbf{z}(t)$ has low rank (rank is 2), which means that at time t, the channels are highly correlated. Figure 2.4 displays reconstruction error as a function of number of factors. It shows that two factors are capable of capturing all the dynamics of the input time series which is reasonable since since the generated time series has no temporal dynamics (they are temporally independent even though they have high spatial correlation).

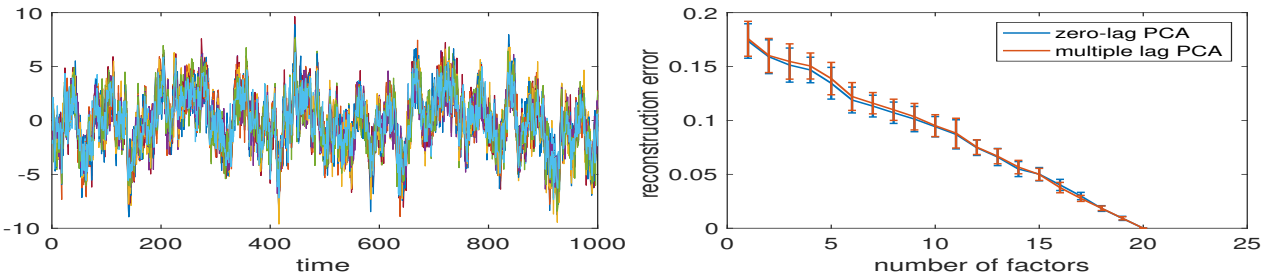


(a) Plot of $\mathbf{z}_{tr}(t)$ (b) Correlation matrix of $\mathbf{z}(t)$ (c) Reconstruction error

Figure 2.4: Simulation result. The time series data have high spatial correlation but low temporal correlation. The reconstruction error drops to zero when the number of factors is greater than 2, for both methods. Two methods have similar reconstruction error evaluated on the test data.

Setting 3: the multivariate time series has high spatial correlation and are temporally correlated

In this set up, we consider the time series $\mathbf{z}(t)$ that has both high spatial correlation and high temporal correlation. That is to say, $\mathbf{z}(t_1)$ and $\mathbf{z}(t_2)$ are correlated and $z_i(t)$ and $z_j(t)$ are also correlated. $\mathbf{z}(t)$ is generated using Equation 2.3.1, where $\mathbf{f}(t)$ follows a first order autoregressive process $f(t) = 0.9f(t-1)+\epsilon(t)$, and $Q(h) = 0$ for all $h \neq 0$. In this setting, all components of the $\mathbf{z}(t)$ are highly correlated and thus we expect that using factors with very low dimension is capable of capturing most of the variation of the original signal. We also expect the two methods have similar performance since the observations are synchronized. The training time series plot and the reconstruction errors are displayed in Figure 2.5. Two methods have similar reconstruction errors. The reconstruction error shows a decreasing trend as the number of factors increases. When the number of factors reaches 20, which is same as the dimension of the observations, the encoding-to-decoding procedure is equivalent to an identity transformation.



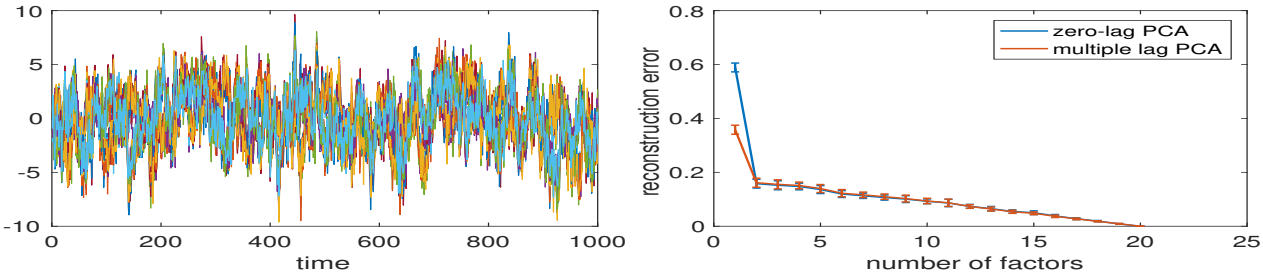
(a) Plot of $\mathbf{z}_{tr}(t)$

(b) Reconstruction error

Figure 2.5: Simulation result. The time series data have strong temporal and spatial correlation.Two methods have similar reconstruction errors.

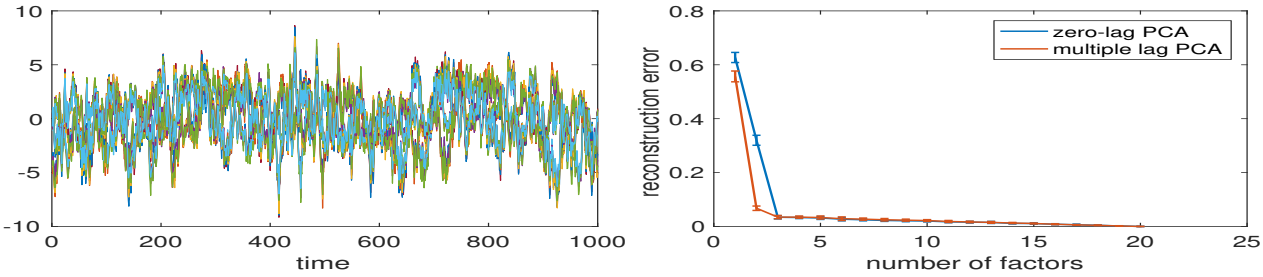
Settings 4 and 5: Components of the multivariate time series are phase-shifted

In this section, we performed simulations where the generated time series has phase-shift. It is important to investigate shifts in time series because it is possible to have lead-lag relationships between EEGs in a region. We consider the autoregressive source process $f(t)$ as described in Section 2.3.1. We generated two types of shifted time series following the description in Table 2.1. Using this setting, we expect that method 2 will have better performance comparing to method 1, as method 2 uses information from all lags, and thus is more capable of capturing the temporal dynamics. It is observable from both Figure 2.6 and 2.7 that (1.) the time series plots show clearly clustered pattern, where within each cluster the signals are more synchronized, (2.) the second model, where the factor is a filtered version of the signal at all time points, gives a lower reconstruction error when the number of factors is smaller than the number of shifted clusters, and (3.) after the number of factors reaches the number of shifted time series clusters, the decreasing rate of the reconstruction error drops dramatically and the two models have similar performance. The decreasing rate of the reconstruction error can be useful in estimating the number of synchronized clusters appeared in the time series data.



(a) Plot of $\mathbf{z}_{tr}(t)$ (b) Reconstruction error

Figure 2.6: Plot of shifted time series (two clusters) (left) and the reconstruction error evaluated on the test data (right). The reconstruction results show that the second model outperforms the first one when number of factors is 1 and that the two models have similar performance when the number of factors is greater than 1.



(a) Plot of $\mathbf{z}_{tr}(t)$ (b) Reconstruction error

Figure 2.7: Plot of shifted time series (three clusters) (left) and the reconstruction error evaluated on the test data (right). The reconstruction results show that the second model outperforms the first one when number of factors is smaller than 3 and that the two models have similar performance when the number of factors is greater than 3.

2.3.2 A systematic comparison of the first factor and the averaged signal

Among the approaches for summarizing signals, the average across channels seems to be the most simple and straightforward way. In this simulation, we compare the power spectrum density of factor 1 and that of the averaged signal. We consider a source $\mathbf{f}(t)$ that follows an AR(2) process defined by $\mathbf{f}(t) = 1.976\mathbf{f}(t-1) - 0.980\mathbf{f}(t-2) + \epsilon(t)$. This process has spectrum concentrated on the 8-12 Hertz frequency band. Then the observations, denoted

by $\mathbf{z}(t)$ are generated via

$$\begin{aligned} \mathbf{s}(t) &= \mathbf{f}(t) + \eta(t), \\ \mathbf{z}(t) &= \sum B_j \mathbf{s}(t-j) + \mathbf{e}(t), \end{aligned}$$

where $\eta(t)$ and $\mathbf{e}(t)$ represent the Gaussian white noise, the loading coefficients B_j satisfy $B_0(k) = 0$ for $k = 11, \ldots, 20$; $B_{40}(k) = 0$ for $k = 1, \ldots, 10$; and $B_j = 0$ for $j \neq 0$ or 40. We repeat the procedure for 100 times and then the pointwise 95% confidence interval is computed. The result is shown in Figure 2.8. It can be seen that in the case where the signals are not well synchronized, the average fails to capture important spectrum features of the original signals.

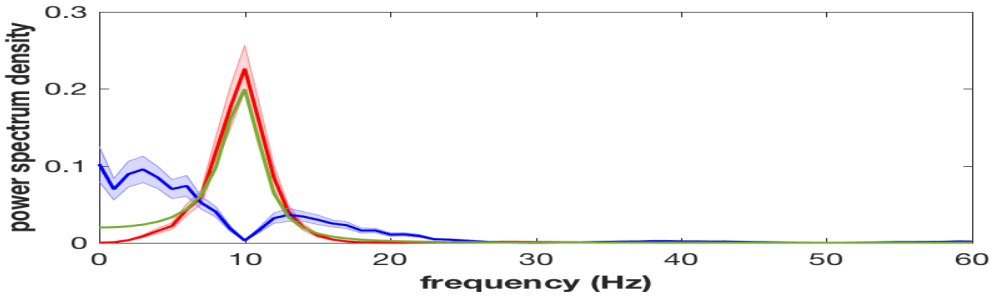


Figure 2.8: Estimated power spectrum density for the top factor (red) and the averaged signal over all 20 channels (blue). The green curve is the true power spectrum density of the $AR(2)$ process that generates $\mathbf{f}(t)$. The shaded area is the point-wise 95% confidence interval for the mean, computed using 100 simulations.

2.3.3 Exploratory analysis of the EEG data

In this section, we perform exploratory analysis on real EEG data. The key challenge in analyzing EEG is the high dimensionality of the data. Computing dependence between regions or channels can be difficult due to the high dimensionality. Our goal here is address the dimensionality problem by deriving signal summaries (factors) of the EEGs in each region

(e.g. SMA, left Pf) and then characterizing the dynamics and connectivity using the factor signals and the encoding/decoding functions.

The data were recorded during a motor learning study performed in the Stroke Rehab laboratory of our collaborator. The dataset contains EEG recordings for multiple subjects, where for each subject, 180 trails of 1 second EEG signals were recorded. The sampling rate of the data is 1000 Hz and number of channels is 256. The raw EEG data have been pre-processed by (1.) applying low pass filter at 50 Hz and (2.) using visual inspection and independent component analysis (ICA) to remove artifacts due to muscle activity, eye blinks and heart rhythms. Various analysis have been performed on the dataset, including using brain connectivity as predictor for ability of motor skill acquisitions (Wu et al., 2014) and the analysis of curves of log periodograms using functional boxplots (Ngo et al., 2015). The goal of the exploratory analysis is to gain better understanding of the EEG data as well as the models that we used.

Computing regional summaries (factors)

It can be observed that the EEG signals are very highly synchronized, which means EEG at channel i is highly correlated with EEG at channel j at the same time t. It also appears that signals within the same region (e.g. SMA and left Pr) have higher correlation, comparing to that of the signals in different regions. Due to these high spatial correlations, it is sensible to represent these EEGs in terms of low dimensional summaries that capture the most variation in these EEG signals. Figure 2.9 shows the reconstruction of EEGs at SMA region using the linear convolution encoder. It can been seen that as the number of factors increases, the magnitude of the residuals decrease. The top two summary signals (factors) computed using EEGs from SMA region and left Pre-frontal region are plotted in Figure 2.10 and the proportion of total variation accounted by theses factors are shown in Figure 2.11. The results for both regions show that factors with very low dimension (less than 3) can represent

most of the variation of the original signals. This is consistent with the fact that the EEGs are highly correlated spatially due to volume conduction.

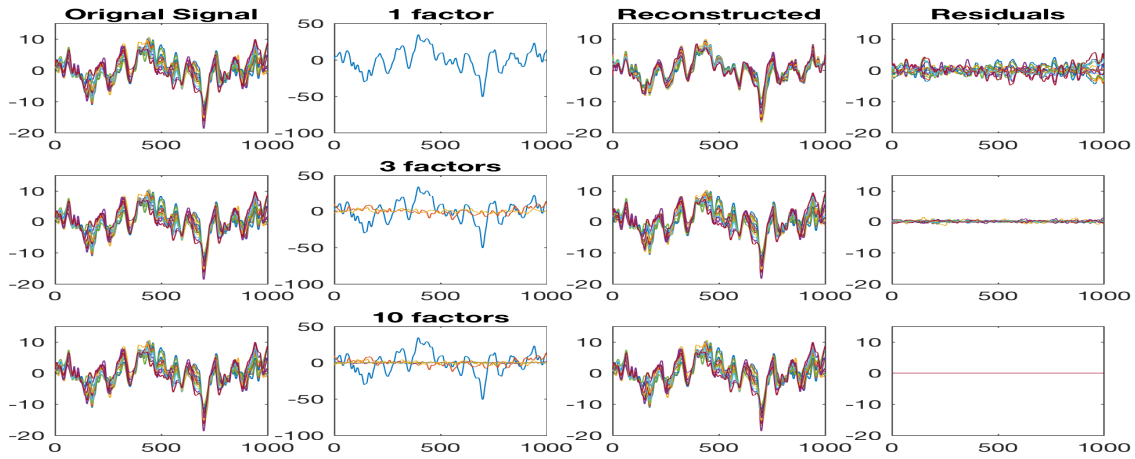


Figure 2.9: Signal compression using spectral PCA (method 2). Column one shows the original EEG time series at SMA region, column two shows the factors computed via spectral PCA, column three shows the reconstructed time series using different number of factors and column four shows the difference between the original signals and the reconstructed signals. Note that as the number of factors increases, the magnitude of the residuals decrease (i.e., the squared error of reconstruction decreases).

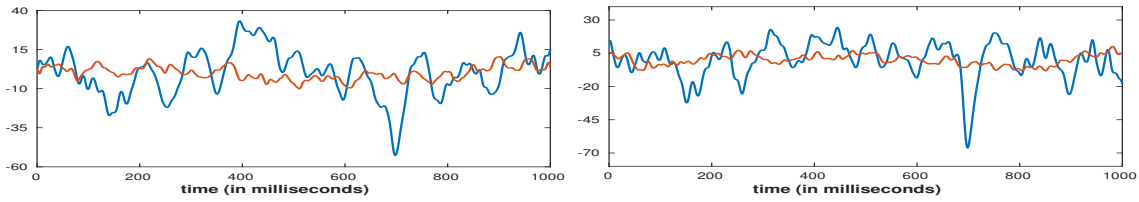


Figure 2.10: Plot of the two factors in the SMA region (left) and the Left Pre-frontal cortex (right) that give the lowest squared reconstruction error.

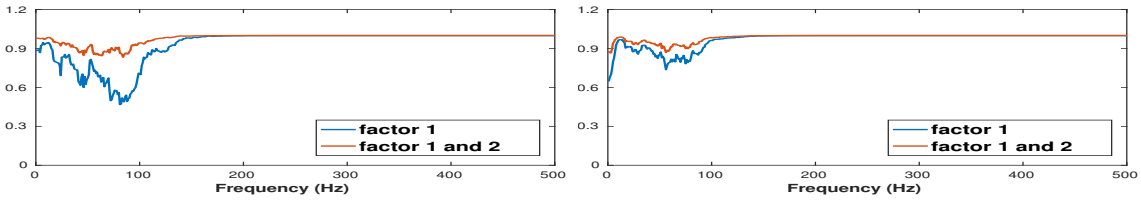


Figure 2.11: Variance accounted by factor 1 only and factors 1 and 2 in the SMA region (left) and the Left Pre-frontal region (right). Note that two factors explain around 90% of the total variance in the high dimensional EEGs.

Properties of the summary factors

Figure 2.12 shows the estimated power spectrum density of the top factors computed for SMA region and left Pf region. It shows that factor 1 in both the SMA and Left Pre-frontal regions capture the alpha oscillations (8-12 Hertz) and low beta (16-30 Hertz). Factor 2 has more power in the delta and theta band oscillations (1-8 Hertz). The power spectrum across 100 EEG epochs are estimated and visualized in Figure 2.13. The results show that the spectrum pattern for the top factors are consistent across trails, where factor 1 concentrates more on alpha oscillations (8-12 Hertz) and factor 2 concentrates more on the delta and theta band oscillations (1-8 Hertz). In order to study the temporal dependence between the factors, we plot the cross-correlation between the top two factors, evaluated for multiple epochs (Figure 2.14). The cross-correlation between factor 1 in SMA region and factor 1 in left Pre-frontal region appears to be very consistent across epochs. The cross-correlation that involves factor 2 also shows some consistent patterns across epochs, although the consistency is weaker comparing to the cross-correlation between factor 1's in two regions.

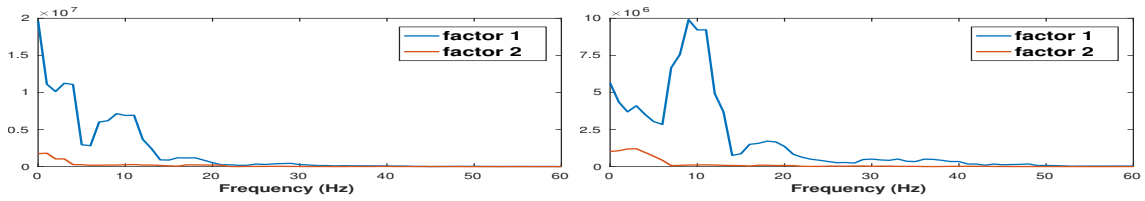


Figure 2.12: Estimated power spectrum for the first two factors in the SMA region (left) and the Left Pre-frontal region (right).

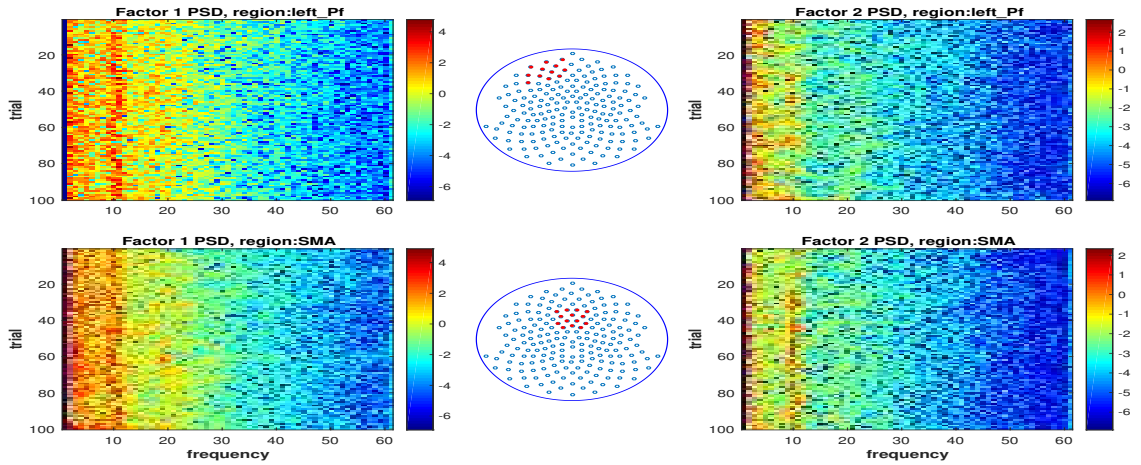


Figure 2.13: Top: SMA region. Estimated power spectrum (log-scale) across 100 epochs of the first factor (left) and the second factor (right). Bottom: Left Pre-frontal region. Estimated power spectrum across 100 epochs of the first factor (left) and the second factor (right).

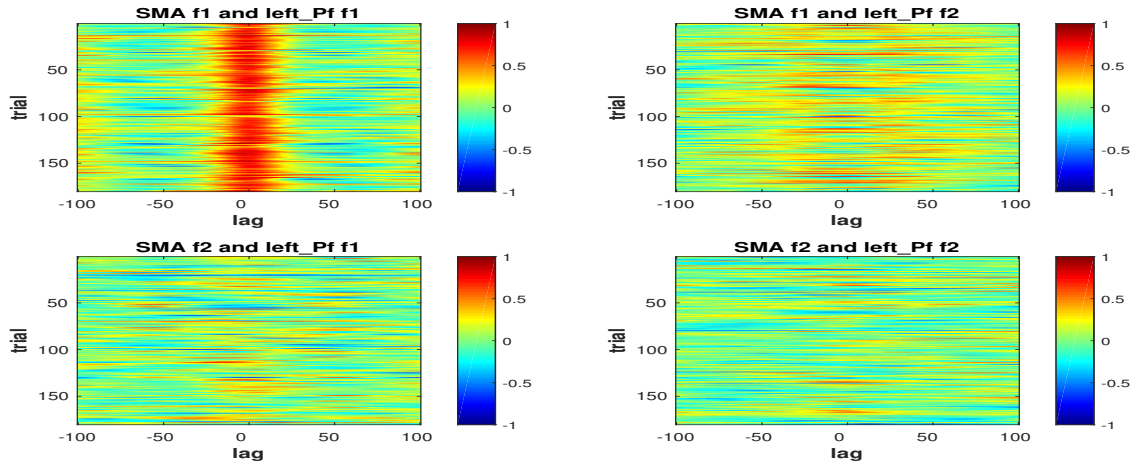


Figure 2.14: Cross correlation between factors in SMA region and Left Pre-frontal region across 180 epochs. The correlations show consistent patterns across epochs.

We also estimated the power spectrum density (PSD) for the factors from 15 subjects, 100 trials for each subject. The PSDs were then averaged over trials. Figure 2.15 displays the mean PSD and the standard deviation across subjects. It appears that the second factor has larger variation in (PSD) across different subjects, comparing to the PSD of the first factor. Both factors have high power around 10 Hz and less.

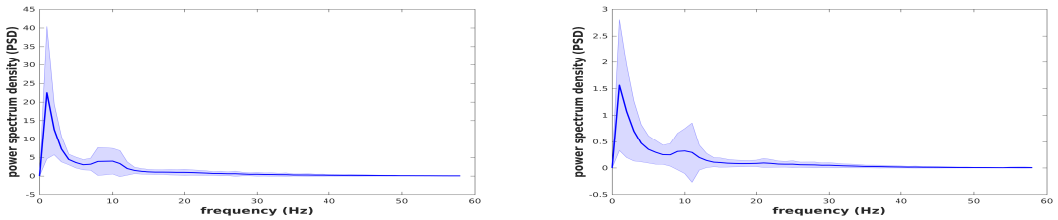


Figure 2.15: Estimated power spectrum for factor 1 (left) and factor 2 (right) in the SMA region, averaged across 15 subject and 100 trials. The shaded area represent the standard error computed from 15 subjects.

2.3.4 Exploratory analysis of fMRI data

The focus of this study is the analysis of EEGs. As a comparison, we also applied the proposed method to fMRI data which has higher dimension. Functional MRI (fMRI) has been widely used in measuring brain activity. Comparing to EEG data, fMRI images have higher spatial resolution ($\sim 10^5$ voxels) but lower temporal resolution. Analyzing the connectivity between ROIs by directly computing the connectivity for each voxel pair is problematic because the number of parameters is a quadratic function of the dimension. We used a resting-state fMRI dataset which is available at NI-TRC (`http://www.nitrc.org/projects/nyu_trt/`). The voxel size is $3 \times 3 \times 3$ and the recording length for each voxel is 197 images with TR/TE=2000/25 milliseconds. The data has been pre-processed with motion correction, normalization, nuisance signal removal and spatial smoothing. The parcellation of the ROIs follows the Anatomical Automatic Labeling (AAL) atlas. In this analysis, we mainly focused on two ROIs: left and right posterior cingulate gyrus (PCC). Specifically, we computed ROI-specific summary signals and then estimated the connectivity between ROIs.

Summary signals

For each ROI, we computed the summary signals using (1.) the top 2 factors via method 1 in Section 2.2.2 and (2.) the averaged signal. We compared the signal summaries and their corresponding spectrum estimates for ROIs left PCC and right PCC (Figure 2.16). For both regions, the first factor accounts for about half of the total variation and the second accounts for about 10%. It also appears that for this dataset, the first factor is very close to the averaged signal. From the estimated power spectrum density, the activity in the fMRI as expected occurs at the lower frequencies as compared to EEGs. It is observed that in both ROIs, the second factor is able to capture additional information on components of higher frequency, not revealed in the spectrum of the averaged signal.

Figure 2.16: Plots of the original time series (top), the summary signals (middle, factors are normalized for easier comparison) and the corresponding spectrum estimated using multitaper method (bottom), for ROIs left PCC (left) and right PCC (right). For left PCC ROI, projection on factor 1 explains 45% of the total variation while the projection on factor 2 explains 10%. For the right PCC ROI, the projection on factor 1 explains 35% of the total variation while the projection on factor 2 explains 14%

Between ROI connectivity

To study the connectivity between ROIs of left PCC and right PCC, we computed the connectivity between their summary signals (i.e., the first and second factor time series). Here use the magnitude of the correlation as a measure of connectivity. The standard approach which is gives the magnitude cross-correlation between the averaged time series gave a connectivity measure of 0.6925. As demonstrated in the simulation studies and in EEG analysis, there are advantages to using factors which are best represent brain activity in each region (in the sense of producing the lowest squared reconstruction error). Connectivity using factors can be characterized in a number ways. By characterizing connectivity between two regions using only the first factors, we obtain a connectivity value of 0.8043. The connectivity from the first factors is higher than that obtained from averaged fMRI time series (0.80 vs 0.69) which could be due in part to the fact that the first factor is optimal at extracting the brain signal in a region. The weights for the first factor are obtained data-adaptively rather than constrained to be equal which is the case for the averaged time series. Note that the connectivity between the second factors is 0.1814 which is capturing another facet of connectivity beyond that explained by the first factor. Another approach to obtain a global measure of connectivity jointly between the multiple factors from different ROIs (e.g. between two factors in the left and right PCC) is to use the RV coefficient ((Escoufier, 1973)). It is a multivariate version of the squared Pearson correlation and it will reduce to the standard squared Pearson correlation when both variables are univariate. Given two random variables $\mathbf{z}_1 \in \mathbb{R}^{n_1}$ and $\mathbf{z}_2 \in \mathbb{R}^{n_2}$, the RV coefficient is computed as $\mathrm{RV}_{\mathbf{z}_1\mathbf{z}_2} = \frac{\mathrm{Tr}[\Sigma_{\mathbf{z}_1\mathbf{z}_2}\Sigma'_{\mathbf{z}_1\mathbf{z}_2}]}{\sqrt{\mathrm{Tr}[\Sigma^2_{\mathbf{z}_1\mathbf{z}_1}]\mathrm{Tr}[\Sigma^2_{\mathbf{z}_2\mathbf{z}_2}]}}$, where $\Sigma_{\mathbf{z}_i\mathbf{z}_j} = \mathbb{E}[\mathbf{z}_i\mathbf{z}'_j]$. Moreover, if we approximate $\mathbf{z}_1$ and $\mathbf{z}_2$ using the factor representation described in Section 2.2.2, we will have $\widehat{\mathbf{z}}_i = B_i\mathbf{f}_i$ for $i = 1, 2$, where B_i is an orthogonal matrix (i.e., $B'_iB_i = I$). Then the RV coefficient between $\mathbf{z}_1$ and $\mathbf{z}_2$ can be approximated using the factor representation as $RV_{\widehat{\mathbf{z}}_1\widehat{\mathbf{z}}_2} = \frac{\mathrm{Tr}[\Sigma_{\mathbf{f}_1\mathbf{f}_2}\Sigma'_{\mathbf{f}_1\mathbf{f}_2}]}{\sqrt{\mathrm{Tr}[\Sigma^2_{\mathbf{f}_1\mathbf{f}_1}]\mathrm{Tr}[\Sigma^2_{\mathbf{f}_2\mathbf{f}_2}]}}$, because $\Sigma_{\widehat{\mathbf{z}}_i\widehat{\mathbf{z}}_j} = B_i\Sigma_{\mathbf{f}_i\mathbf{f}_j}B'_j$ and thus $\mathrm{Tr}[\Sigma_{\mathbf{z}_1\mathbf{z}_2}\Sigma'_{\widehat{\mathbf{z}}_1\widehat{\mathbf{z}}_2}] = \mathrm{Tr}[B_i\Sigma_{\mathbf{f}_i\mathbf{f}_j}B'_jB_j\Sigma'_{\mathbf{f}_i\mathbf{f}_j}B'_i] = \mathrm{Tr}[B'_iB_i\Sigma_{\mathbf{f}_i\mathbf{f}_j}\Sigma'_{\mathbf{f}_i\mathbf{f}_j}] = \mathrm{Tr}[\Sigma_{\mathbf{f}_i\mathbf{f}_j}\Sigma'_{\mathbf{f}_i\mathbf{f}_j}]$. $RV_{\widehat{\mathbf{z}}_1\widehat{\mathbf{z}}_2}$

is equivalent to the weighted average of the squared correlations between $\mathbf{f}_1$ and $\mathbf{f}_2$ where the weight is proportional to the product of the corresponding eigenvalues. Based on the summary factors, the estimated RV coefficient between signals of left and right PCC ROIs is 0.6003 (0.7748^2).

2.4 Conclusion and future work

In this work, we developed exploratory procedures for analyzing high dimensional brain signals under the presence of high multi-collinearity by using low-dimensional factor representations for pacellated brain regions. We considered two schemes for estimating the factors: the conventional time-domain PCA and the spectral-domain PCA, which respectively assumes the factors as an instantaneous mixing and a linear filtering of the high-dimensional measured signals. Compared to averaged signals typically used in ROI-based analysis of fMRI data, the factors are able to characterize the variability and dependence across voxels, are derived from a parametric modeling beyond simply a dimension reduction method. We evaluated (benchmark) the performance of the proposed methods via numerical experiments by applying the models on simulated time series with different spatial and temporal correlation structures. Besides, the spectral PCA is more advantageous than the conventional PCA, in capturing the high-frequency oscillatory brain activities as measured by the EEGs, and accounts for both the temporal and cross-correlation between regions. To the best of our knowledge, we are among the limited studies to apply spectral PCA to analyze EEG and fMRI data. We performed exploratory analysis on a motor-task EEG and a resting-state fMRI data to assess the suitability of our method as compared to the traditional averaging approach. The results for both of the simulation and exploratory analysis show that learning low-dimensional representations (factors) has potential benefits for subsequent modeling of the connectivity in high dimensional time series, because the factors are capable of capturing

the dynamics of the data (i.e., temporal dynamics, variation) and the detailed spectral structure while reducing dimension (complexity) of the original problem. We also implemented the methods in a Matlab toolbox with graphical interface that allows users to interactively explore, process and analyze the data in a convenient way. Currently we only explored linear projection or convolution in summarizing region-specific signals, where the solution can be computed explicitly using eigenvalue decomposition. It may lack the ability to characterize non-linear association. Our future work includes a comprehensive evaluation of the methods. For example, we will evaluate the ability of the models in capturing the temporal dynamics and at the same time, quantify the artifact that might be induced by the mixing. Moreover, to make the package more comprehensive, we shall include other emerging measures of dependence such as isolated coherence (Pascual-Marqui et al., 2014; Ombao and Van Bellegem, 2008; Fiecas and Ombao, 2011; Wang et al., 2016a) and other more general (possibly non-linear) methods for obtaining summary signals (Peña and Yohai, 2015).

Chapter 3

Modeling Effective Connectivity in High-Dimensional Cortical Source Signals

To study effective connectivity among sources in a densely voxelated (high-dimensional) cortical surface, we develop the source-space factor VAR model. The first step in our procedure is to estimate cortical activity from multichannel electroencephalograms (EEG) using anatomically constrained brain imaging methods. Following parcellation of the cortical surface into disjoint regions of interest (ROIs), latent factors within each ROI are computed using principal component analysis. These factors are ROI-specific low-rank approximations (or representations) which allow for efficient estimation of connectivity in the high-dimensional cortical source space. The second step is to model effective connectivity between ROIs by fitting a VAR model jointly on all the latent processes. Measures of cortical connectivity, in particular partial directed coherence, are formulated using the VAR parameters. We illustrate the proposed model to investigate connectivity and interactions between cortical ROIs during rest. This chapter is a reprint with permission from Wang et al. (2016b).

3.1 Introduction

Functional connectivity of a brain network, i.e., dependence between distinct brain regions of interest (ROIs) in a network, has provided important insights on various brain functions and abnormalities. Effective connectivity, which is a more specific measure of cross-dependence, quantifies the *directed* causal influence of one neuronal region over another Friston et al. (1994). Effective and functional connectivity are both usually inferred from brain signals such as functional magnetic resonance imaging (fMRI) and multichannel electroencephalogram (EEG) He et al. (2011). These two modalities measure different facets of brain activity: hemodynamic by fMRI and electrical by EEG. While EEG has high temporal resolution (in the millisecond scale), it suffers from limited spatial resolution. Nevertheless, EEG has been widely used for studying brain activity for a number of good reasons: it captures fast temporal dynamics in brain activity; it is non-invasive; it is relatively easy and inexpensive to collect. Examples on the utility of EEG for studying brain function and detecting neurological disorders are described in Grill-Spector et al. (2006); Schwartz et al. (2010); Wu et al. (2014); Nunez and Srinivasan (2006b). However, the EEG signal captured by a specific sensor on the scalp is not a direct measurement of activity of a single source (local subpopulation of neurons) but rather the summation of synchronous activity of many sources (many populations of neurons) over the entire cortex that are similarly oriented perpendicular to the local cortical surface. Therefore, the EEG signals can be regarded as a mixture of underlying cortical source signals plus some noise. This mixing effect may introduce spurious instantaneous correlations between EEG signals which may lead to misinterpretation of connectivity estimates Nunez et al. (1997).

In this chapter, we consider this challenge of estimating effective connectivity on the (high-dimensional) source space by developing a novel statistical model with rigorous estimation method. Our novel procedure uses multichannel EEG to model and estimate effective connectivity between regions on the brain cortical surface where the number of sources, as modeled

by dipoles distributed on a densely voxelated cortical surface, is very large. The common approach employs a two-stage procedure Ding et al. (2007); Babiloni et al. (2005); Astolfi et al. (2005, 2007b, 2008); Hui et al. (2010): Stage 1: Source reconstruction from scalp EEGs via the inversion of the mixing process, and Stage 2: Fitting the VAR model to the estimated source signals. Estimates of various connectivity quantities such as Granger causality Granger (1969b), partial directed coherence (PDC) Baccalá and Sameshima (2001b) and directed transfer function (DTF) Kaminski and Blinowska (1991) are then derived from the VAR model. Our proposed method also follows the two-step approach but differs at Stage 2. We do not directly fit a VAR model to the estimated source due to the high dimensionality of the source space. To overcome this problem, our model takes the additional step by summarizing source activity at a cortical region via a factor time series whose dimension is lower than that of the source signals. Cross-regional dependency is then characterized through the VAR model on the factors, and PDC is used as a frequency-domain measure of effective connectivity between the ROIs. Compared to other approaches that summarize ROI-specific activity by averaging signals Astolfi et al. (2007a), our model is more comprehensive because it is able to characterize connectivity at both dipole level and region level.

Two classes of source reconstruction approaches have been developed. The first is the dipole modeling technique which depends on the physical head forward model which assumes that the EEG sources are equivalent dipole currents located on the cortical surface. However, solving the ill-conditioned inverse problem poses a major challenge, which we address in this chapter by identifying the optimal solution that minimizes the reconstruction error subject to the regularization of the spatial smoothness of the source. The other class is the blind source separation technique such as independent component analysis (ICA) Makeig et al. (1996b); Onton et al. (2006) by assuming statistically independent sources. However, both the source signals and un-mixing matrix need to be estimated and the current approaches are based on the unrealistic assumptions including statistical independence and non-Gaussianity of the sources and dimension of the source space to be no greater than the dimension of the

observation space, to avoid the identifiability problem. A new family of integrated approaches has recently been introduced. These combine linear mixing model of the sources to EEGs and standard VAR model for the causal interactions between sources. The variants of this single-step approach, which involve joint estimation of the mixing matrix and the VAR source connectivity, include the convolutive ICA (CICA) Haufe et al. (2010) and the VAR+ICA model formulated in the state-space form Gómez-Herrero et al. (2008); Chiang et al. (2012); Cheung et al. (2010); Chiang et al. (2009). However, these do not take into consideration the high-dimensionality of the source space, where the number of source signals are larger or comparable to the number of time points.

In the first class of approach, the distributed source model assumes that each point on the cortical grid has electrical signal that contributes to the scalp-recorded EEG Michel et al. (2004). For a meshgrid of cortex area with 3 mm spacing, the number of grid points (equivalent dipoles) can be greater than 10^4, which might be comparable to the number of time points of the signal. This renders estimation and inference of the standard VAR modeling for such high-dimensional source connectivity difficult, even for a moderately large network, due to a large number of parameters. It is not feasible to fit a VAR model directly on the high-dimensional source signals, since the dimension of the parameter space is in the order of square of the dimension of the signals. The conventional least squares estimators of the VAR parameters could be inconsistent. This leads to unreliable estimators for the subsequently constructed frequency-based measures of directed connectivity between cortex regions. One plausible solution proposed in Davis et al. (0) is to impose sparsity on the VAR parameters. Conventional technique involves partitioning the source space into smaller number of cortical regions of interest (ROIs) and extracting a mean source signal for each ROI by averaging the magnitudes of all dipole currents within the ROI Babiloni et al. (2005); Haufe et al. (2013). This approach that uses within-ROI-averaged-source as the factor, while sensible, is ad-hoc and cannot handle multi-scale connectivity. It is useful only for global (between-ROI) connectivity but not local (between-dipole) connectivity. Moreover,

summarizing activity via simple averaging is not optimal because it assigns, a priori, equal weights on all sources within an ROI. It implies loss of information on the degree of variability of source activity with a ROI.

We adopt a model-based dimensionality-reduction approach using factor analysis to estimate the large-scale effective cortical connectivity. We develop a multi-layer factor-analyzed VAR model for the high-dimensional source signals estimated from EEGs, that allows simultaneously for (1) reliable and computationally-efficient estimation of directed dependence between huge number of cortical sources, and (2) multi-scale analysis of hierarchical, modular connectivity at both regional (within cortical ROI) and global (between cortical ROIs) level. The proposed model generalizes the factor-VAR model developed for fMRI in our recent work Ting et al. (2014) by partitioning the massive spatial dimension into lower-dimensional sub-models. More precisely, the entire brain source space is first anatomically-parcellated into a set of many ROIs. Moreover, to further reduce the dimensionality, we use a factor model to represent each cortical ROI, where the source signals within each region are characterized (summarized) by a small number of latent factors. This is based on the assumption that low-rank (highly-correlated) source data within a region lies on a subspace that is of lower dimension than the source space. We apply principal component analysis (PCA) to the source signals (reconstructed via standardized low resolution brain electromagnetic tomography (sLORETA)Pascual-Marqui et al. (2002)) within each region to estimate the region-specific latent factors and the factor loadings. PCA can data-adaptively identify a subspace of a lower dimension. It preserves most of the dipole-wise variation of the localized source activity explained by the few principal components. This source activity cannot be captured by taking the regional average signal Zhou et al. (2009). In contrast with many other EEG-based cortical connectivity studies, we apply the PCA for dimension-reduction on the source signals instead of on the EEGs, as a pre-modeling step Haufe et al. (2014).

To establish directed dependence between ROIs, the temporal dependency structure of com-

bined factors over all regions will be characterized by a VAR model. Since the factors are of much lower dimension compared to the original source space, it would be possible to fit a VAR model with reasonably good accuracy. The estimated VAR parameters are then used to construct the directed and non-directed coherence for both local (dipole) level and global (between ROI) level.

The remainder of the chapter is organized as follows. Section II and III describe our proposed multi-scale factor VAR model and the estimation procedure for analyzing high-dimensional EEG source connectivity. Section IV presents the evaluation results on simulated source signals generated using VAR parameters derived from reconstructed EEG sources. Section V presents the evaluation results on a real resting-state EEG data set, followed by a conclusion in Section VI.

3.2 Source-Space Factor VAR Model for Cortical Connectivity Analysis

In this section, we develop the source-space factor VAR model which is a novel approach to modeling effective connectivity in high-dimensional cortical source space. The approach partitions the large cortical space into disjoint anatomical regions and then derives region-specific factors via principal component analysis. Next, we fit a VAR model to the latent factors which have lower dimension than the original source space. Thus, effective connectivity in high-dimensional space can be efficiently estimated. Finally, the proposed model gives a multi-scale analysis of both intra-region (between dipoles in a ROI) and inter-region (between ROIs) *directed* cortical connectivity.

3.2.1 EEG Signal Model

According to biophysical models, scalp EEG signals can be modeled as a linear mixing of sources on the cortex. Let $\mathbf{Z}(t) = [Z_1(t), \ldots, Z_N(t)]' \in \mathbb{R}^N, t = 1, \ldots, T$ be the electrical signals of length T due to neuronal activity at N dipoles on the voxelated cortical surface. The observed EEG signals $\mathbf{X}(t) \in \mathbb{R}^M$ can be represented as

$$\mathbf{X}(t) = \mathbb{A}\mathbf{Z}(t) + \mathbf{N}(t) \tag{3.2.1}$$

where $\mathbf{N}(t)$ is the measurement and physiological noise; $\mathbb{A} \in \mathbb{R}^{M \times N}$ is the mixing matrix (which is also known as the lead field matrix). The above model for EEG (3.2.1) has been widely used in EEG source reconstruction and studies on EEG source connectivity Hassan et al. (2014); Grech et al. (2008); Baillet et al. (2001b).

3.2.2 Cortical Source Models

Factor Model for a Cortical Region

The entire cortex area can be parcellated into a finite number of disjoint regions of interest (ROIs) according to the anatomical structure of the brain. Let R be the total number of ROIs. We denote by $\mathbf{Z}_r(t) \in \mathbb{R}^{n_r}$ the activities of n_r dipoles at each region $r \in \{1, 2, \ldots, R\}$ ($\sum_{r=1}^{R} n_r = N$). To achieve further dimension reduction, we summarize activity in each cortical region via a set of common m_r latent factors. Specifically, we assume the source activity $\mathbf{Z}_r(t)$ at region r to be driven by factor activity $\mathbf{f}_r(t) \in \mathbb{R}^{m_r}$ with loading matrix $\mathbf{Q}_r$, where the dimension of $\mathbf{f}_r(t)$ is much lower than that of $\mathbf{Z}_r(t)$, (i.e. $m_r \ll n_r$). The factor model at region r is given by

$$\mathbf{Z}_r(t) = \mathbf{Q}_r \mathbf{f}_r(t) \tag{3.2.2}$$

where $\mathbf{Q}_r \in \mathbb{R}^{n_r \times m_r}$ is the factor loading matrix for region r and $\mathbf{f}_r(t)$ represents factor activity at time t. For identifiability or unique model decomposition, we assume the following constraints on both $\mathbf{Q}_r$ and $\mathbf{f}_r(t)$: $\mathbf{Q}_r'\mathbf{Q}_r = \mathbf{I}_{m_r}$ and $\mathbb{C}\text{ov}(\mathbf{f}_r(t))$ is a diagonal matrix with distinct positive diagonal elements.

Factor Model for the Entire Cortex

The source activity of the whole cortex area denoted by $\mathbf{Z}(t) = [\mathbf{Z}_1(t)', \mathbf{Z}_2(t)', \ldots, \mathbf{Z}_R(t)']'$ can be represented by a global factor model

$$\mathbf{Z}(t) = \mathbf{Q}\mathbf{f}(t) \tag{3.2.3}$$

where $\mathbf{f}(t) = [\mathbf{f}_1(t)', \mathbf{f}_2(t)', \ldots, \mathbf{f}_R(t)']' \in \mathbb{R}^M$ and $\mathbf{Q} = \text{diag}\{\mathbf{Q}_1, \mathbf{Q}_2, \ldots, \mathbf{Q}_R\}$ are, respectively, the concatenated global $M \times 1$ factor signal and $N \times M$ loading matrix over all regions, with $M = \sum_{r=1}^{R} m_r$. The factors within a cortical region are uncorrelated but the factors between different regions can be correlated or dependent. Note that the uncorrelatedness within a region only happens at an instantaneous time (lag zero), however, allows for correlations at lags greater than zero. This lagged dependence will be modeled using an VAR process as follows.

VAR Model for Factor Activity

According to the factor model (3.2.3), $\mathbf{Z}(t)$ is the instantaneous mixing of the factor activity $\mathbf{f}(t)$. Therefore the temporal inter-dependence structure for $\mathbf{Z}(t)$ can be characterized by the temporal dependence structure of factor $\mathbf{f}(t)$. We use a VAR model with order P (denoted

VAR(P)) to characterize the temporal dependence structure in $\mathbf{f}(t)$

$$\mathbf{f}(t) = \sum_{\ell=1}^{P} \mathbf{\Phi}^{\mathrm{f}}(\ell)\mathbf{f}(t-\ell) + \boldsymbol{\eta}(t) \tag{3.2.4}$$

where $\mathbf{\Phi}^{\mathrm{f}}(\ell) \in \mathbb{R}^{M\times M}$ represents the VAR coefficient matrix at lag ℓ with block structure

$$\mathbf{\Phi}^{\mathrm{f}}(\ell) = \begin{pmatrix} \mathbf{\Phi}_{\mathbf{f}_1\mathbf{f}_1}(\ell) & \dots & \mathbf{\Phi}_{\mathbf{f}_1\mathbf{f}_R}(\ell) \\ \vdots & \ddots & \vdots \\ \mathbf{\Phi}_{\mathbf{f}_R\mathbf{f}_1}(\ell) & \dots & \mathbf{\Phi}_{\mathbf{f}_R\mathbf{f}_R}(\ell) \end{pmatrix}.$$

The diagonal block $\mathbf{\Phi}_{\mathbf{f}_r\mathbf{f}_r}(\ell)$ summarizes the lagged auto-correlation within each region, while the off-diagonal block $\mathbf{\Phi}_{\mathbf{f}_j\mathbf{f}_k}(\ell)$, where $j \neq k$, captures the cross-dependence between regions j and k; and $\boldsymbol{\eta}(t) \in \mathbb{R}^M$ represents the Gaussian white noise with mean zeros and covariance matrix $\mathbf{\Sigma}_\eta$.

VAR Model for Source Activity

Note that the dimension of the parameter space for a VAR model grows quadratically with the dimension of the signal. Therefore, a VAR model for an $N-$ dimensional signal will have $P \times N^2$ parameters that could be much greater than the total number of observations NT. In the EEG source study, given the high-dimensionality of $\mathbf{Z}(t)$, it would not be feasible to fit a VAR model via ordinary least-squares (LS) without some regularization.

In our procedure, we construct a high-dimensional VAR model with a unique hierarchical, block structure for the source space, based on the lower dimensional factor VAR model in (3.2.4). We use a VAR model to characterize the temporal inter-dependence structure for the factor activity $\mathbf{f}(t)$, as described in Section 3.2.2. We show that the temporal inter-dependence structure for the high-dimensional source activity $\mathbf{Z}(t)$ can be characterized by a VAR model with a special structure, which is derived from the parameter of the VAR

model for the factors, with the following steps.

At this point, we shall assume that $\mathbf{Z}(t) = \mathbf{Q}\mathbf{f}(t)$ (i.e., the significant activity in the sources are captured by the factors). Substituting the VAR factor process (3.2.4) in the model defined in (3.2.3) gives

$$\begin{aligned}
\mathbf{Z}(t) &= \mathbf{Q}\Big(\sum_{\ell=1}^{P} \mathbf{\Phi}^{\mathrm{f}}(\ell)\mathbf{f}(t-\ell) + \boldsymbol{\eta}(t)\Big) && (3.2.5)\\
&= \sum_{\ell=1}^{P} \mathbf{Q}\mathbf{\Phi}^{\mathrm{f}}(\ell)\mathbf{Q}'\mathbf{Q}\mathbf{f}(t-\ell) + \mathbf{Q}\boldsymbol{\eta}(t) && (3.2.6)\\
&= \sum_{\ell=1}^{P} \mathbf{Q}\mathbf{\Phi}^{\mathrm{f}}(\ell)\mathbf{Q}'\mathbf{Z}(t-\ell) + \mathbf{Q}\boldsymbol{\eta}(t). && (3.2.7)
\end{aligned}$$

Finally, we have the following factor VAR model for $\mathbf{Z}(t)$

$$\mathbf{Z}(t) = \sum_{\ell=1}^{P} \mathbf{\Phi}^{\mathrm{Z}}(\ell)\mathbf{Z}(t-\ell) + \mathbf{E}(t) \qquad (3.2.8)$$

where $\mathbf{\Phi}^{\mathrm{Z}}(\ell) = \mathbf{Q}\mathbf{\Phi}^{\mathrm{f}}(\ell)\mathbf{Q}' \in \mathbb{R}^{N\times N}$ is the global VAR coefficient matrix at lag ℓ for the entire cortex, projected from a much lower-dimensional matrix $\mathbf{\Phi}^{\mathrm{f}}(\ell)$. This is a block matrix where the diagonals $\mathbf{\Phi}_{\mathbf{Z}_r\mathbf{Z}_r}(\ell) = \mathbf{Q}_r\mathbf{\Phi}_{\mathbf{f}_r\mathbf{f}_r}(\ell)\mathbf{Q}_r'$ and the off-diagonal block $\mathbf{\Phi}_{\mathbf{Z}_j\mathbf{Z}_k}(\ell) = \mathbf{Q}_j\mathbf{\Phi}_{\mathbf{f}_j\mathbf{f}_k}(\ell)\mathbf{Q}_k'$ capture, respectively, the inter-dipole effective source connectivity within a region and across different regions. $\mathbf{E}(t) = \mathbf{Q}\boldsymbol{\eta}(t) \in \mathbb{R}^N$ is a Gaussian white noise process with zero mean and covariance matrix $\mathbf{\Sigma}_E = \mathbf{Q}\mathbf{\Sigma}_\eta\mathbf{Q}'$.

3.2.3 Measures of Source Connectivity

In this section, we develop measures of brain source connectivity at different levels of organization: local (between dipoles in an ROI) and global (between ROIs). We use partial

directed coherence (PDC) to quantify the directed connectivity between ROIs and between dipoles within an ROI.

Inter-dipole Effective Connectivity

PDC has been introduced to infer frequency-specific effective connectivity between dipoles. The inter-dipole PDC can be interpreted as the direct impact of a change in the amplitude of an oscillatory activity (specifically at frequency ω) in one dipole on the amplitude of oscillatory activity in another dipole (accounting for the effects of the oscillatory activity in other dipoles). It can be treated as frequency-domain analogue of Granger causality. PDC between the high-dimensional dipole sources can be constructed from our proposed factor VAR model.

Let $\mathbf{\Phi}^{\mathrm{Z}}(\omega) = \mathbf{I} - \sum_{\ell=1}^{P} \mathbf{\Phi}^{\mathrm{Z}}(\ell) \exp(-i2\pi\omega\ell/\Omega_s)$ be the Fourier transform of the VAR coefficient matrix $\mathbf{\Phi}^{\mathrm{Z}} = [\mathbf{\Phi}_{ij}(\omega)]_{1\leq i,j\leq N} \in \mathbb{R}^{N\times N}$ at frequency ω, where $\mathbf{I}$ is the identity matrix, and Ω_s is the sampling frequency. The PDC derived from the VAR model (3.2.8) is defined as

$$\pi_{ij}(\omega) = \frac{|\mathbf{\Phi}^{\mathrm{Z}}_{ij}(\omega)|}{\sqrt{\sum_{k=1}^{N} \mathbf{\Phi}^{\mathrm{Z}}_{kj}(\omega)(\mathbf{\Phi}^{\mathrm{Z}}_{kj}(\omega))^*}}. \tag{3.2.9}$$

Here, $\pi_{ij}(\omega) \in [0, 1]$. It gives an indication of the strength of linear directed influence of ω-oscillatory activity at the j-th dipole, denoted $Z_j(t)$, on the i-th dipole $Z_i(t)$, relative to the total influence of $Z_j(t)$ on all dipoles. A value close to one indicates that the causal influences originating from the dipole j on cortex are directed, for the most part, toward the dipole i. A value of zero indicates no directed influences from dipole j to i. The matrix $\boldsymbol{\pi}(\omega) = [\pi_{ij}(\omega)]_{1\leq i,j\leq N} \in \mathbb{R}^{N\times N}$ characterizes a network of directed interactions between dipoles of the entire cortex at frequency ω.

Inter-dipole Functional Connectivity

In cases where a measure of functional connectivity is desired we derive coherence from the VAR model as follows. The spectral matrix of $\mathbf{Z}(t)$ at frequency ω is $\mathbf{S}^{\mathrm{Z}}(\omega) = \mathbf{Q}\mathbf{S}^{\mathrm{f}}(\omega)\mathbf{Q}'$, where $\mathbf{S}^{\mathrm{f}}(\omega)$ is the spectral matrix of $\mathbf{f}(t)$, which can be computed as

$$\mathbf{S}^{\mathrm{f}}(\omega) = \mathbf{H}^{\mathrm{f}}(\omega)\mathbf{\Sigma}_{\eta}(\mathbf{H}^{\mathrm{f}}(\omega))^{*} \tag{3.2.10}$$

where $\mathbf{H}^{\mathrm{f}}(\omega) = (\mathbf{\Phi}^{\mathrm{f}}(\omega))^{-1}$, $\mathbf{\Phi}^{\mathrm{f}}(\omega) = \mathbf{I} - \sum_{\ell=1}^{P} \mathbf{\Phi}^{\mathrm{f}}(\ell)\exp(-i2\pi\omega\ell/\Omega_s)$. The coherence between dipoles i and j is then defined as

$$\rho_{ij}^{2}(\omega) = \frac{|\mathbf{S}_{ij}^{\mathrm{Z}}(\omega)|^{2}}{\mathbf{S}_{ii}^{\mathrm{Z}}(\omega)\mathbf{S}_{jj}^{\mathrm{Z}}(\omega)}. \tag{3.2.11}$$

Inter-ROI Effective Connectivity

To characterize connectivity between a pair of ROIs, we use the average value of pairwise connectivity between dipole pairs at the two regions. Specifically, given regions r_1 and r_2, the inter-region connectivity between the two, denoted by $C_{r_1r_2}(\omega)$ is computed as

$$C_{r_1r_2}(\omega) = \frac{1}{|U_{r_1}|\,|U_{r_2}|} \sum_{i\in U_{r_1}, j\in U_{r_2}} \pi_{ij}(\omega) \tag{3.2.12}$$

where U_{r_1} and U_{r_2} are the sets of dipoles within regions r_1 and r_2 respectively, $|U_r|$ represents the cardinality of the set U_r therefore we have $|U_r| = n_r$. Alternatively, the connectivity between multivariate time series in frequency domain can be characterized using general coherence Pascual-Marqui (2007a) or frequency decomposition of Granger causality Geweke (1982).

3.3 Model Estimation and Inference

We develop a two-step approach to estimate connectivity in high-dimensional source space. Compared to the conventional VAR model (without the lower-rank representation), our proposed approach has lower estimation error. Moreover, our approach is less computationally intensive since the factor space usually has a much lower dimension. In Step 1, source activity on a densely voxelated cortical space is estimated. In Step 2, factors within each cortical ROI are computed and then concatenated. A VAR model is fit to the concatenated factors. From the estimated VAR parameters, connectivity between cortical ROIs and dipoles are formulated and estimated.

3.3.1 Cortical Source Reconstruction

The primary sources of EEG are believed to be intracellular currents in dendritic trunks of the pyramidal neuron in cerebral cortex. The cortex sources are not directly measurable and hence need to be estimated from the observable EEG signals. It is an ill-conditioned inversion problem when the number of EEG sensors (usually in the order of 10^2) is smaller than the dimension of the cortex source space (which can be of order 10^4).

Obtaining the Source Mixing Matrix

The lead field matrix $\mathbb{A}$ in the EEG observation model (3.2.1) can be computed by applying the boundary element method (BEM) on a discretized realistic head MRI template, as shown in Fig. 3.1. In this work, we use the Colin 27 MRI head template Holmes et al. (1998).

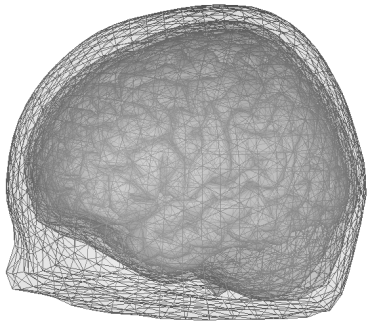


Figure 3.1: BEM head model computed from the Colin 27 head template. The BEM head model is computed using the software package OpenMEEG(compute) Gramfort et al. (2010); Kybic et al. (2005), and the discretized head is visualized using the software Brainstorm Tadel et al. (2011)

Inverse Source Reconstruction from EEG

Given the lead field matrix $\mathbb{A}$ and EEG observations $\mathbf{X}(t)$, the cortical sources $\mathbf{Z}(t)$ are reconstructed using standardized low resolution brain electromagnetic tomography (sLORETA) method as described in Pascual-Marqui et al. (2002). It is a brain imaging method that computes brain activity from EEG measurements. It provides an instantaneous linear solution that has zero localization error under ideal conditions. The sLORETA method starts from a solution of minimum norm estimate (MNE) Hämäläinen and Ilmoniemi (1994b) method that minimizes the objective function of (3.3.1) with respect to $\mathbf{Z}(t)$ and c, given lead field matrix $\mathbb{A}$, EEG observation $\mathbf{X}$ and α.

$$F = \sum_{t=1}^{T} (\|\mathbf{X}(t) - \mathbb{A}\mathbf{Z}(t) - c\mathbf{1}\|^2 + \alpha\|\mathbf{Z}(t)\|^2) \tag{3.3.1}$$

The solution of the minimization gives the estimate $\widehat{\mathbf{Z}}(t) = \mathbf{G}\mathbf{X}(t)$, where $\mathbf{G} = \mathbb{A}'\mathbf{H}(\mathbf{H}\mathbb{A}\mathbb{A}'\mathbf{H} + \alpha\mathbf{H})^{-1}$ with $\mathbf{H} = \mathbf{I} - \mathbf{1}\mathbf{1}'/\mathbf{1}'\mathbf{1}$. The regularization parameter α is determined by cross-validation, which is the same method as described in Pascual-Marqui et al. (2002). The sLORETA estimation for the source is then obtained by standardizing $\widehat{\mathbf{Z}}(t)$ using its variance estimate.

3.3.2 Parameter Estimation

Principal component analysis (PCA), a commonly used dimension-reduction technique, projects the original high-dimensional signal to a space of lower dimension while preserving as much of the variation in the original signal. We apply PCA to the estimated regional source $\widehat{\mathbf{Z}}_r(t)$ to estimate the regional factor activity $\mathbf{f}_r(t)$ and its loading $\mathbf{Q}_r(t)$ in (3.2.2). Then we concatenate these regional factor estimates to form the estimate for the global factor model in (3.2.3) for the entire cortex. A VAR model for the factor space in (3.2.4) is fitted on the estimated global factor time series, and then projected to the VAR model in source space in (3.2.8) using the factor loadings. Various directed source connectivity quantities in the frequency domain for both the dipole and regional level, are then constructed from the source-space VAR parameters.

Let $\widehat{\mathbf{Z}}_r = [\widehat{\mathbf{Z}}_r(1), \ldots, \widehat{\mathbf{Z}}_r(T)]' \in \mathbb{R}^{T \times n_r}$ be the reconstructed cortical sources for region r. Here are the estimation steps.

- Compute the sample covariance matrix of $\mathbf{Z}_r(t)$, denoted as $\widehat{\Sigma}_{Z_r Z_r}$:

$$\widehat{\Sigma}_{Z_r Z_r} = \widehat{\mathbf{Z}}_r' \widehat{\mathbf{Z}}_r / T = \frac{1}{T} \sum_{t=1}^{T} \widehat{\mathbf{Z}}_r(t) \widehat{\mathbf{Z}}_r'(t) \tag{3.3.2}$$

  where $\widehat{\mathbf{Z}}_r(t)$ has zero mean.

- Estimate the factor loading matrix $\mathbf{Q}_r \in \mathbb{R}^{n_r \times m_r}$ and the factors $\mathbf{f}_r(t)$ based on the eigenvalue-eigenvector decomposition of $\widehat{\Sigma}_{Z_r Z_r}$. Let $\lambda_1 \ldots \lambda_{n_r}$ be the unique eigenvalues of $\widehat{\Sigma}_{Z_r Z_r}$, in decreasing order and let $\mathbf{V}_1, \ldots, \mathbf{V}_{n_r} \in \mathbb{R}^{n_r}$ be the corresponding orthonormal eigenvectors. The estimator of $\mathbf{Q}_r$ can be defined as

$$\widehat{\mathbf{Q}}_r = [\mathbf{V}_1, \ldots, \mathbf{V}_{m_r}] \tag{3.3.3}$$

where m_r is the dimension of the regional factor activity $\mathbf{f}_r(t)$ which can be determined to be the smallest dimension that exceeds some prespecified amount of region-specific variation explained by the factors.

- Compute the regional factor activity $\mathbf{f}_r \in \mathbb{R}^{T \times m_r}$

$$\widehat{\mathbf{f}}_r = \widehat{\mathbf{Z}}_r \widehat{\mathbf{Q}}_r \tag{3.3.4}$$

where $\widehat{\mathbf{f}}_r = [\widehat{\mathbf{f}}_r(1), \ldots, \widehat{\mathbf{f}}_r(T)]'$.

- Compute the estimate of the VAR coefficient $\widehat{\Phi}^{\mathrm{f}}(\ell)$ for factor activity $\mathbf{f}(t)$ by least-squares fitting on $\widehat{\mathbf{f}}_r$. The optimal VAR order $\widehat{P}$ is the minimizer of the Akaike information criterion (AIC):

$$\mathrm{AIC}(P) = \log |\widehat{\mathbf{\Sigma}}_\eta(P)| + \frac{2}{T} P M^2 \tag{3.3.5}$$

where $\widehat{\mathbf{\Sigma}}_\eta(P) = T^{-1} \sum_{t=1}^{T} \widehat{\boldsymbol{\eta}}(t) \widehat{\boldsymbol{\eta}}(t)'$ is the residual covariance matrix without a degree of freedom correction, where $\widehat{\boldsymbol{\eta}}(t) = \widehat{\mathbf{f}}(t) - \sum_{\ell=1}^{P} \widehat{\Phi}^{\mathrm{f}}(\ell) \widehat{\mathbf{f}}(t-\ell)$.

- Compute the VAR coefficient estimate for source activity $\mathbf{Z}(t)$ by substitution $\widehat{\mathbf{\Phi}}^{\mathrm{Z}}(\ell) = \widehat{\mathbf{Q}} \widehat{\mathbf{\Phi}}^{\mathrm{f}}(\ell) \widehat{\mathbf{Q}}'$ where $\widehat{\mathbf{Q}} = \mathrm{Diag}\{\widehat{\mathbf{Q}}_1, \ldots, \widehat{\mathbf{Q}}_R\}$.

- Estimate the *between-dipole source connectivity* by computing the PDC $\widehat{\pi}_{ij}(\omega)$ from dipole j to dipole i via substitution of $\widehat{\mathbf{\Phi}}^{\mathrm{Z}}(\ell)$ in (3.2.9).

- Estimate the *between-ROI source connectivity* $\widehat{C}_{r_1 r_2}(\omega)$ by summarizing the estimated inter-dipole PDC $\widehat{\pi}_{ij}(\omega)$ over the region from r_2 to r_1 according to (3.2.12).

3.4 Numerical Experiments

In this section, we evaluate the performance of our proposed factor-analytic VAR model for estimating high-dimensional cortical effective connectivity. In our simulation, we study the role of the length of the time series (relative to the dimension of the parameters) under the setting where the VAR parameters are derived cortical sources estimated from the actual EEG data using a realistic head model. Thus, the simulation setting was made to be as realistic as possible by using EEG data-derived parameters. We benchmark our proposed factor-VAR model-based estimator with the conventional least squares (LS) estimator and the ridge estimator commonly used for large-dimensional VAR.

3.4.1 Data Generation

In order to generate time series that emulate dependence structure as the cortical sources underlying EEGs, we first reconstructed the source signals, using sLORETA, from resting-state EEG signals (This dataset was provided by Dr S. C. Cramer, Neurology, UC Irvine). We used source signals $\mathbf{Z}(t)$ from a set of $N = 120$ dipoles from 6 cortical regions (20 dipoles per region, regions include LPF, RPF, LF, RF, LC and RC). The length of the EEG recording used in the above source reconstruction was 1000 samples (corresponding to one trial).

Next we fitted the factor VAR model of (3.2.8) on $\mathbf{Z}(t)$ to obtain the VAR coefficients matrices $\mathbf{\Phi}^{\mathrm{Z}}$ and the covariance matrix of residuals $\mathbf{\Sigma}_E$. The order of the VAR model was selected by AIC with maximum lag equals to 4. We generated the simulated source time series according to the model (3.2.8). The $\mathbf{\Phi}^{\mathrm{Z}}$ was assumed known in this simulation study and used as ground-truth to compare different estimators. We repeated the simulation for 100 times.

3.4.2 Model Evaluation

We evaluated the accuracy of the model in estimating the directed connectivity in terms of PDC in the source space. Let $\widetilde{\boldsymbol{\pi}}(\omega)$ be the connectivity matrix whose elements are given in $\pi_{ij}(\omega)$ (as defined in (3.2.9)) and $\widehat{\boldsymbol{\pi}}(\omega)$ be the estimated connectivity matrix, the error of the estimation is evaluated using the Frobenius norm of the difference between the two. More specifically, the estimation error denoted by Λ, is computed as follows

$$\Lambda(\Omega) = \frac{1}{|\Omega|} \sum_{\omega \in \Omega} \frac{\|\widehat{\boldsymbol{\pi}}(\omega) - \widetilde{\boldsymbol{\pi}}(\omega)\|_F}{\|\widetilde{\boldsymbol{\pi}}(\omega)\|_F} \tag{3.4.1}$$

where Ω is the set of frequencies of interest, here we used $\Omega = \{\omega = 0, \ldots, 50\}$, which covers delta (0-4 Hz), theta (4-8 Hz), alpha (8-16 Hz), beta (16-32 Hz) and gamma (32-50 Hz) frequency bands, and $\|\mathbf{H}\|_F = \mathrm{tr}\,(\mathbf{H}'\mathbf{H})^{1/2}$ denotes the Frobenius norm of matrix $\mathbf{H}$.

3.4.3 Results

We compared the performance of our factor VAR model-based estimator with the traditional ordinary LS VAR estimator and the L_2-regularized or ridge estimator, in recovering the ground-truth connectivity from simulated signals. The ridge estimator which uses the L_2 norm penalty on the large-dimensional vector of AR coefficients in the LS regression, is better conditioned than the LS for estimation of large VAR models. The regularization parameter was set $\lambda = 0.1$, as suggested by Korobilis (2013). We investigated the impact of time series dimension on the estimation performance. We increased the length of the simulated time series T with the fixed dimension $N = 120$, to create different scenarios of dimensionality via the ratio of $d = T/N$ with $d < 1$ $(T < N)$, $d = 1$ $(T = N)$ and $d > 1$ $(T > N)$. Fig. 3.2 shows the comparison of the estimation errors for the different methods, for increasing time series lengths relative to the dimension. The error bars were computed

as the standard error of the estimation error over the 100 simulations. When T is smaller or comparable to N, the factor-VAR estimator clearly outperforms both the LS and ridge estimators, with substantially lower mean-squared estimation errors and standard errors. This implies the proposed VAR estimator has an improved accuracy and consistency under the high-dimensional small sample size settings. Nevertheless, the ridge estimator which imposes a shrinkage prior on the large-dimensional VAR coefficients, slightly outperformed the LS estimator. As expected, when the sample size T increases, the Frobenius norm error decreases for all estimators. However, the proposed factor VAR model produced the fastest rate of decline in the mean squared error. In addition, when T is sufficiently large, the LS estimator is comparable to the factor-VAR and the ridge methods.

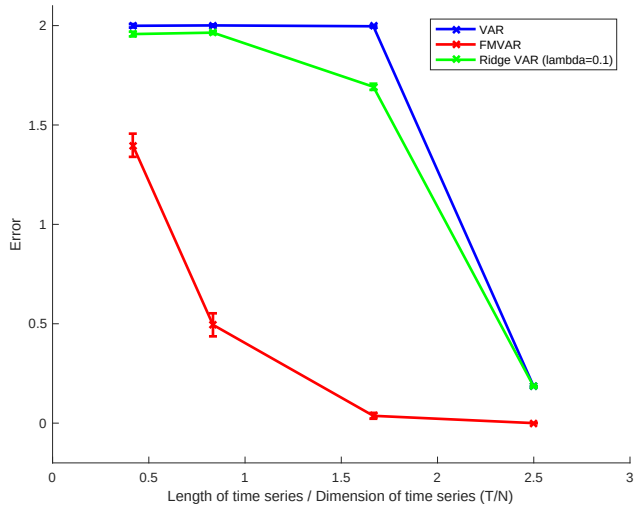


Figure 3.2: Comparison of estimation error of the effective source connectivity by the least-squares, ridge and factor-VAR estimator for increasing length of time series relative to dimension T/N.

3.5 Application to Resting-State EEG

In this section, we reconstructed the cortical source activity from EEG recordings obtained for a subject during resting-state. We then studied both effective connectivity between dipoles *within* each cortical ROI and the effective connectivity *between* cortical ROIs by

fitting a factor VAR model on the reconstructed sources. Our model provides a more reliable and computationally efficient tool for analyzing hierarchical, directed connectivity structure in high-dimensional source signals.

3.5.1 Experimental Data & Cortical Parcellation

The original EEG data has 256 channels, of which 194 artifact-free channels have been selected for analysis. The sampling rate was 1000 Hz but we only performed analysis from the delta to the beta band (0-32 Hz). We analyzed 60 EEG epochs, each with $T = 3000$ observations (3 seconds).

The cortex area was first parcellated into 148 sub-regions according to the Destrieux atlas using software package FreeSurfer Fischl et al. (2004). The Destrieux atlas was then grouped to 14 cortex regions using software package BrainstormTadel et al. (2011), according to the brain anatomy. The 14 cortical ROIs, as shown in Fig. 3.3, include left/right prefrontal (LPF, RPF), frontal (LF, RF), central (LC, RC), parietal (LP, RP), temporal (LT, RT), occipital (LO, RO) and limbic (LL, RL). The number of dipoles on the cortex used is $N = 2800$, where each cortex region contains 200 dipoles. This poses a typical problem of high-dimensional estimation with finite samples.

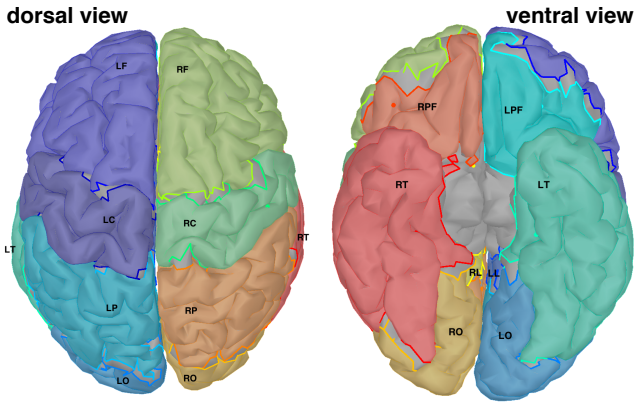


Figure 3.3: Cortical regions: left/right prefrontal (LPF, RPF), frontal (LF, RF), central (LC, RC), parietal (LP, RP), temporal (LT, RT), occipital (LO, RO) and limbic (LL, RL).

3.5.2 Resting-state Cortical Connectivity

While resting-state brain connectivity networks have been extensively investigated in fMRI studies Smith et al. (2013), most studies on EEG focus on task-based source connectivity (e.g., motor-imagery for brain-computer interface Haufe et al. (2014); and visual task Cheung et al. (2010)). There are few studies on the resting-state brain networks based on EEG, e.g. in Chen et al. (2008); Knyazev et al. (2011) that analyzed the default mode network at different frequency bands. However, these are in the scalp-EEG sensor-space rather than the source-space thus posing challenges in interpretations due to volume conduction. To the best of our knowledge, our study is among the first to report results on large-scale resting-state cortical connectivity from EEG sources, at different hierarchical scales along with the additional measures of directionality.

We reconstructed cortex source activity from the observed EEG using the procedure described in Section 3.3.1. Then we applied the proposed factor model to estimate the connectivity in source space at the following two scales (1.) local: between-dipole effective connectivity; and (2.) global: between-ROI effective connectivity.

Effective Connectivity Between Dipoles

We characterize the between-dipole effective connectivity via PDC that is estimated from our factor VAR model. The estimation results for the four conventional frequency bands of interest are shown in Fig. 3.4. The estimates using only small number of latent factors are able to reveal the presence of complex interactions between a large number of dipoles (in this case there were $n_r = 200$ dipoles for *each* of the $R = 14$ ROIs, with the entire source space dimension $N = 2800$). The number of latent factors used for each region is shown in Table 3.1. The number of factors is determined such that they explain at least 95% of the variation of the signal within the region. The total number of factors underlying the

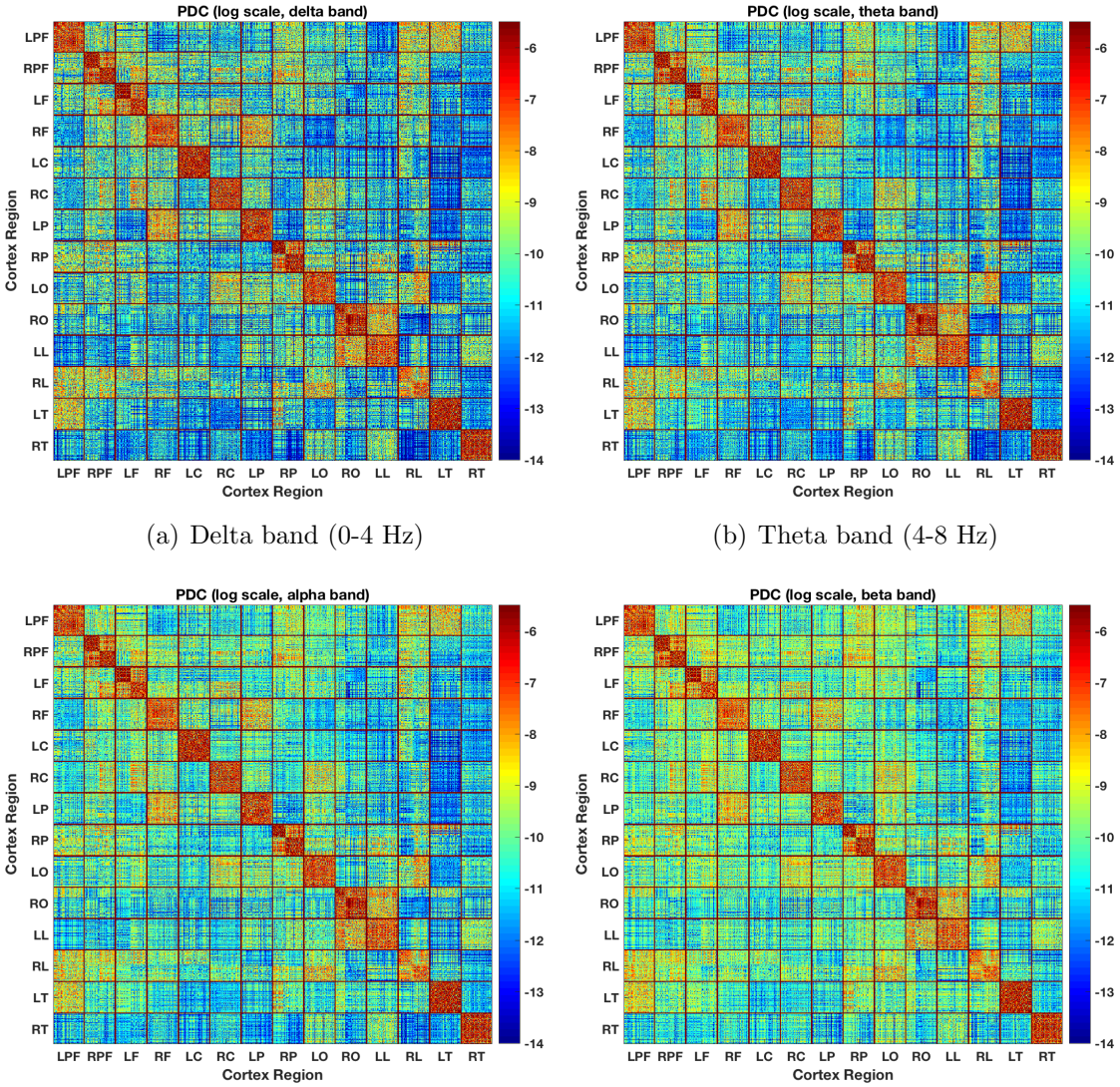
entire cortex only range from $M = 45$ to $M = 49$ across different epochs, which achieved substantial dimension reduction, $M << N$.

Our method identified the modular organization of brain network during resting-state, where dipoles within an ROI are more densely and strongly connected compared to that between ROIs. Note that stronger effective connectivity is indicated by the intense-red color for the connectivity blocks along the diagonals and weaker connections by the less-intense color on the off-diagonals. This suggests that spatial proximity could play a role in the effective connectivity. That is, directed dependence between sub-populations of neurons (whose activity is summarized by a dipole) appears to be stronger when these sub-populations in closer proximity that between sub-populations that are far away. This phenomenon is prevalent across all frequency bands although there appears to be more spatially widespread between-ROI PDC at the higher frequency bands (beta (16-32 Hz)), compared the lower frequency bands, namely, delta (0-4 Hz), theta (4-8 Hz) and alpha (8-16 Hz). We note though that this ought to be further investigated. Quite naturally there are challenges to extracting results from high-dimensional time series but the proposed model could be a useful tool for this purpose.

The results from our analysis on effective connectivity is consistent with findings in other fMRI studies of brain networksMeunier et al. (2010, 2009). This points to the ability of the proposed method to reveal this modular brain structure based on the cortical sources and thus provides a measure for the directionality of the connections between dipoles on the cortex.

Effective Connectivity Between Regions

Fig. 3.5 shows the estimated between-ROI PDC, averaged over results from 60 epochs (each epoch is a recording for 3 seconds). The first goal in this analysis was to determine the *average* PDC in the default mode network (DMN) over the entire resting state of 180



(a) Delta band (0-4 Hz)

(b) Theta band (4-8 Hz)

(c) Alpha band (8-16 Hz)

(d) Beta band (16-32 Hz)

Figure 3.4: Estimated inter-dipole effective connectivity measured by PDC for different frequency bands, averaged across 60 epochs. For convenience of visualization, the value used in the heatmap is computed as $\log(\widehat{\pi}_{ij})$, where $i, j = 1, \ldots, N$ with $N = 2800$

seconds.

The results suggest asymmetry between inter-region forward and backward causal flow. Stronger directed connectivity occurred at the higher frequency bands such as the alpha and beta rhythms, compared to lower frequencies. The figure clearly displays pronounced bi-directional directed influences or information flows between the limbic, prefrontal, frontal

Table 3.1: The min and max number of factors used for each region for the factor model when applying on 60 time series segments. The number of factors was determined such that at least 95% of the total variation was accounted.

Number of Factors (m_r) for each Region					
Region	m_r	Region	m_r	Region	m_r
LPF	3-3	LC	3-3	LO	3-3
RPF	4-5	RC	3-3	RO	3-4
LF	4-5	LP	3-4	LL	3-3
RF	3-4	RP	4-5	RL	4-4
LT	3-3	RT	2-3		

and parietal regions (e.g., RL-LPF, RL-RPF, RL-LF and RP-RPF and LP-RF), which belongs to the DMN *et al.* (2001), one of the well-known resting-state networks (RSNs). We can also see that the limbic regions are strongly inter-connected with many other brain regions. Our method based on cortical sources has identified this dense connectivity of the limbic region during the resting-state, in agreement with other findings based on fMRI. This is the location of the posterior cingulate cortex (PCC) which is the major hub of the DMN Fransson and Marrelec (2008). We also detected strong connectivity between regions of the attentional networks which include the frontal-parietal sub-network *et al.* (2008) (e.g. RF-LP) and the temporal-parietal junction areas (e.g. LT-RP). The estimates show increased activation and correlated neuronal activities in these large-scale resting-state networks, as reported in many fMRI studies van den Heuvel and Hulshoff (2010); *et al.* (2011).

The similarity of resting-state connectivity pattern as detected here in the EEG source activity with that in the BOLD fluctuations is illuminated by several studies on the relationship between spontaneous EEG oscillations and BOLD activity during rest Mantini et al. (2007); Britz et al. (2010). These studies demonstrated significant correlations between band-limited EEG power spectra and the fMRI BOLD signals in brain regions partly associated with some specific resting-state networks, albeit with rather mixed results. Some studies reported EEG correlate of the DMN activity Mantini et al. (2007) while Britz et al. (2010) did not find any. Moreover, these studies did not directly quantify the association between the between-region

functional connectivity and the EEG power. Although our study does not investigate the BOLD/EEG correlation, we found enhanced effective connectivity at areas of the DMN and the attentional networks, based on reconstructed dipole sources which are less confounded by volume conduction as the scalp-EEGs. Thus, the results of the analysis using our proposed factor VAR model add to the current findings showing the electrophysiological signatures of human-brain resting-state networks based on scalp-EEG or independent components, besides the hemodynamic signature traditionally revealed by fMRI.

Time-Evolving Connectivity Across Epochs

Previous studies on the EEG source connectivity analysis assume temporal stationarity over the entire time course of recordings Haufe et al. (2010); Chiang et al. (2012); Cheung et al. (2010). The second goal in the analysis is to investigate the dynamics of brain effective connectivity over the entire resting state recording of 180 seconds. In Fig. 3.6, we reshaped (vectorized) the $R \times R$ PDC connectivity matrix to a $R^2 \times 1$ column vector with the connectivity C_{ij} located at row $R*(j-1)+i$. Column i represents the connectivity estimated from epoch $i = 1, \ldots, 60$. This allowed a visualization of the dynamics of PDC across the entire resting state period. We noted that there are two distinct patterns on the dynamics of PDC over resting state. The first cluster consists of the following ROIs: LPF, RL and LT; and the second cluster consists of all other 11 regions. Indeed, this cross-section clustering reflects that in Fig. 3.5. It suggests the temporal regions (e.g. LT) has particularly low correlation with other regions for the entire time course, compared to other resting-state connectivity. In addition to spatial clustering, we also wanted to examine if the resting-state connectivity was temporally stationary. Under stationarity, PDC should have remained constant across the entire resting-state. What we observe, however, suggests "local" behavior of stationarity in the effective connectivity structure. The PDC between cortical sources tend to congregate into distinct quasi-stable states/regimes or 'microstates', which remains constant within a

short-time regime but with abrupt transitions across different regimes. This may imply a rapid switching between distinct functional brain networks in the resting-state. In particular, the results show PDC estimates over, e.g. epochs 1-8; epochs 12-18 and epochs 22 -27 are approximately constant. However, since PDC evolved across resting-state, one should be careful to not simply assume stationarity of cortical signals during resting-state. In contrast with the recent microstate analysis of resting-state EEG based on scalp-topographic maps Khanna et al. (2015), we analyzed state-related changes in the connectivity maps directly, between underlying cortical sources.

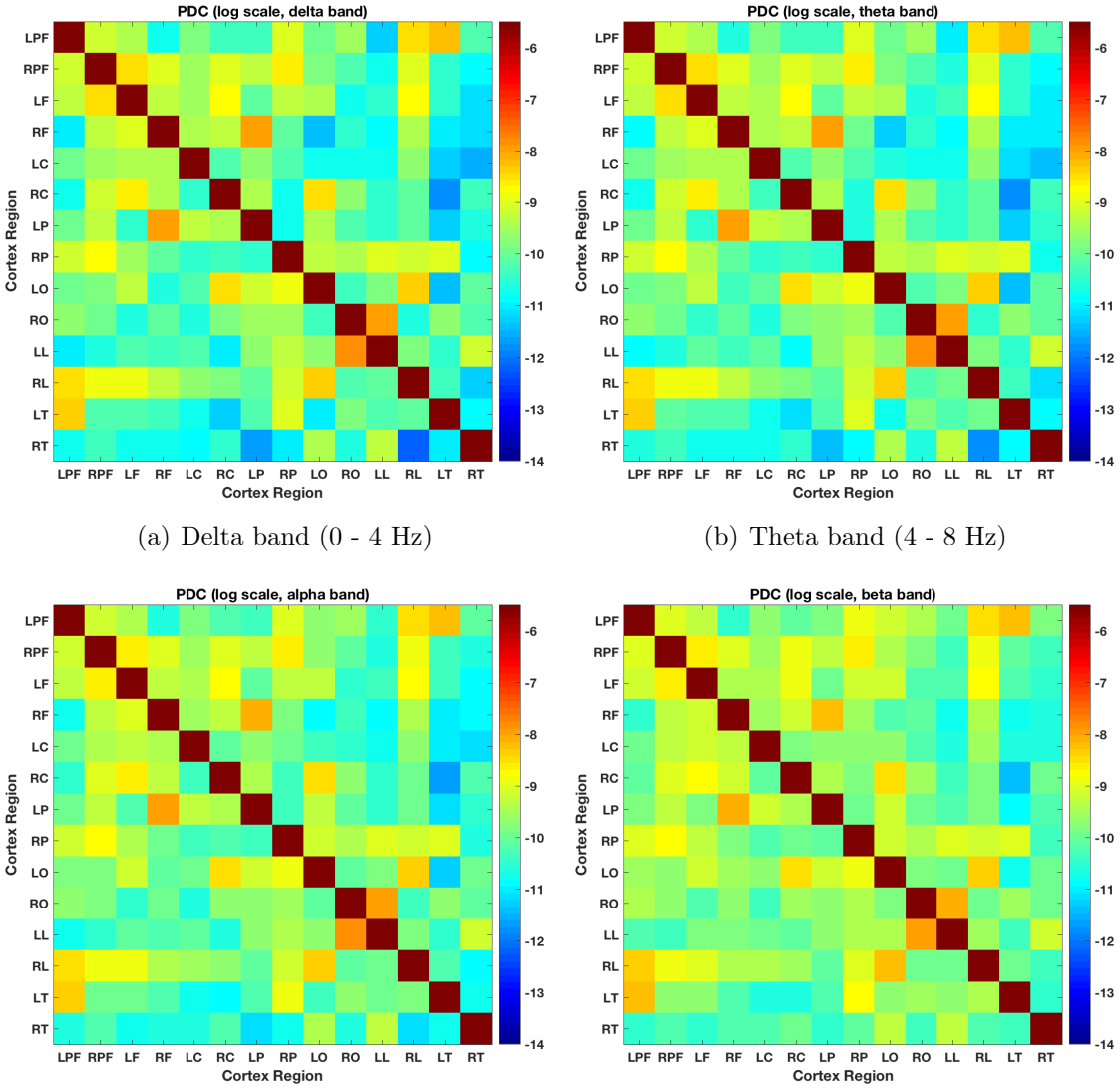
3.6 Conclusion

We developed a procedure for analyzing effective connectivity between high-dimensional dipole sources from a dense grid on the cortical surface. Our method, based on factor analysis, first extracts a small number of factors or summaries of neuronal activity within each cortical ROI. The rationale here is that different cortical sources within a ROI may share common factors as each source is a mixture of these factors. From this commonality we derive the connectivity between the sources. These factors, concatenated from different ROIs, are then modeled as a VAR process. The proposed model provides a powerful tool for dimensionality reduction of high-dimensional source signals.

The proposed factor-VAR model also provides a framework for analyzing effective connectivity of high-dimensional times series at multiple levels or scales. At the dipole level in the source space, the parametric estimate for the factor model can be back-projected to the VAR model for the source space, which provides very powerful tools for subsequent analysis, including effective connectivity such as partial directed coherence; directed transfer function and Granger causality; and functional connectivity such as coherence and partial coherence. At the regional level, effective connectivity can be summarized either by aggregating effec-

tive connectivity between all pairs of dipoles from the regions or directly from the VAR parameters of the factor activities.

Future work involves incorporating the phase-slope index (PSI) Nolte et al. (2008) in the proposed framework. This is promising because PSI is non-parametric and hence potentially flexible in capturing directionality. Moreover, motivated by other improved versions of PDC used in estimating connectivity between the high-dimensional EEG sensors (rather than the sources), we will investigate generalized PDC (gPDC) Luiz A. Baccala (2014) and generalized orthogonalized PDC (gOPDC) Omidvarnia et al. (2014), which are considered to be more robust to noise mixing. The current work was illustrated only to a single-subject data but the proposed framework will be generalized so that it can be used to test for differences in effective connectivity between different experimental conditions; and various populations (e.g., healthy controls vs. patients). Finally, we shall adopt the mixed effects vector autoregressive (ME-VAR) model by Gorrostieta et al. (2012) to account for the variation in effective connectivity among different subjects even in a homogeneous group. Finally, as noted, PDC appears to have evolved during resting-state which could indicate non-stationarity of the cortical sources even during resting-state. One approach will be to adapt the novel model developed in Fiecas and Ombao (2016); Samdin et al. (2016) for evolving brain processes to the high-dimensional setting.



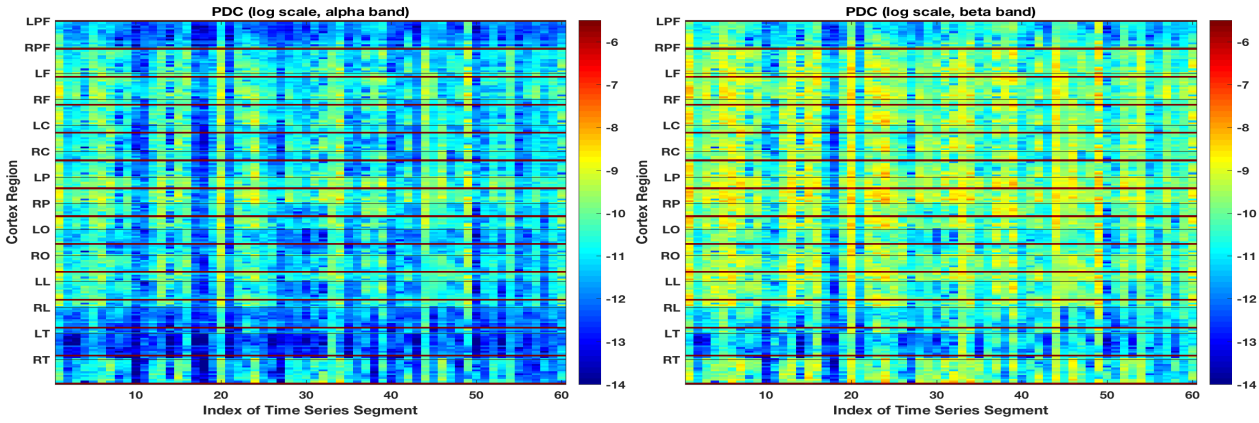
(a) Delta band (0 - 4 Hz)

(b) Theta band (4 - 8 Hz)

(c) Alpha band (8 - 16 Hz)

(d) Beta band (16 - 32 Hz)

Figure 3.5: Band-limited inter-region effective connectivity (in log-scale) summarized from the inter-dipole PDC blocks in Fig. 3.4.



(a) Alpha band (8 - 16 Hz)

(b) Beta band (16 - 32 Hz)

Figure 3.6: Epoch-specific estimates of inter-region PDC at the alpha band (8 - 16 Hz) and the beta band (16 - 32 Hz), plotted across the 60 epochs each with a duration of 3s. In order to visualize the time-variation of connectivity across epochs, we reshaped (vectorized) the $R \times R$ connectivity matrix to a $R^2 \times 1$ column vector with the connectivity C_{ij} located at row $R * (j-1) + i$. We use thin horizontal lines to separate the regions, that is the line is placed after each C_{Rj}, for $j \in \{1, \ldots, R\}$. Column i represents the connectivity estimated from epoch $i = 1, \ldots, 60$.

Chapter 4

Summary and Future Work

4.1 Summary

In summary, the contributions in the dissertation include (1.) we explored the approaches for effectively reducing the dimensionality of high-dimensional time series while preserving the lead-lag relationships; (2.) we implemented a Matlab toolbox with interactive interface that help users to perform exploratory analysis of multivariate time series and (3) based on the low dimensional embeddings, we proposed a factor based vector autoregressive (VAR) model, which is capable of characterizing the lead-lag relationship in high-dimensional brain cortical sources.

In chapter 2, we developed exploratory procedures for analyzing high dimensional brain signals under the presence of high multi-collinearity by using low-dimensional factor representations for pacellated brain regions. We considered two schemes for estimating the factors: the conventional time-domain PCA and the spectral-domain PCA, which respectively assumes the factors as an instantaneous mixing and a linear filtering of the high-dimensional measured signals. Compared to averaged signals typically used in ROI-based analysis of fMRI

data, the factors are able to characterize the variability and dependence across voxels, are derived from a parametric modeling beyond simply a dimension reduction method. We evaluated (benchmark) the performance of the proposed methods via numerical experiments by applying the models on simulated time series with different spatial and temporal correlation structures. The simulation results also show that the spectral PCA is more advantageous than the conventional PCA, in capturing the high-frequency oscillatory brain activities as measured by the EEGs, and accounts for both the temporal and cross-correlation between regions. To the best of our knowledge, we are among the limited studies to apply spectral PCA to analyze EEG and fMRI data. We performed exploratory analysis on a motor-task EEG and a resting-state fMRI data to assess the suitability of our method as compared to the traditional averaging approach. The results for both of the simulation and exploratory analysis show that learning low-dimensional representations (factors) has potential benefits for subsequent modeling of the connectivity in high dimensional time series, because the factors are capable of capturing the dynamics of the data (i.e., temporal dynamics, variation) and the detailed spectral structure while reducing dimension (complexity) of the original problem. We also implemented the methods in a Matlab toolbox with graphical interface that allows users to interactively explore, process and analyze the data in a convenient way.

In chapter 3, we developed a procedure for analyzing effective connectivity between high-dimensional dipole sources from a dense grid on the cortical surface. Our method, based on factor analysis, first extracts a small number of factors or summaries of neuronal activity within each cortical ROI. The rationale here is that different cortical sources within a ROI may share common factors as each source is a mixture of these factors. From this commonality we derive the connectivity between the sources. These factors, concatenated from different ROIs, are then modeled as a VAR process. The proposed model provides a powerful tool for dimensionality reduction of high-dimensional source signals. The proposed factor-VAR model also provides a framework for analyzing effective connectivity of high-dimensional times series at multiple levels or scales. At the dipole level in the source space, the parametric

estimate for the factor model can be back-projected to the VAR model for the source space, which provides very powerful tools for subsequent analysis, including effective connectivity such as partial directed coherence; directed transfer function and Granger causality; and functional connectivity such as coherence and partial coherence. At the regional level, effective connectivity can be summarized either by aggregating effective connectivity between all pairs of dipoles from the regions or directly from the VAR parameters of the factor activities.

4.2 Future Directions

For the low dimensional embedding approaches of time series, we only explored linear projection or convolution in summarizing region-specific signals, where the solution can be computed explicitly using eigenvalue decomposition. It may lack the ability to characterize non-linear association. Our future work includes a comprehensive evaluation of the methods. For example, we will evaluate the ability of the models in capturing the temporal dynamics and at the same time, quantify the artifact that might be induced by the mixing. Moreover, to make the package more comprehensive, we shall include other emerging measures of dependence such as isolated coherence (Pascual-Marqui et al., 2014; Ombao and Van Bellegem, 2008; Fiecas and Ombao, 2011; Wang et al., 2016a) and other more general (possibly non-linear) methods for obtaining summary signals (Peña and Yohai, 2015).

In chapter 3, we proposed a parametric model (VAR) for the effective connectivity in cortical sources. Future work involves incorporating the phase-slope index (PSI) Nolte et al. (2008) in the proposed framework. This is promising because PSI is non-parametric and hence potentially flexible in capturing directionality. Moreover, motivated by other improved versions of PDC used in estimating connectivity between the high-dimensional EEG sensors (rather than the sources), we will investigate generalized PDC (gPDC) Luiz A. Baccala (2014) and generalized orthogonalized PDC (gOPDC) Omidvarnia et al. (2014), which are

considered to be more robust to noise mixing. The current work was illustrated only to a single-subject data but the proposed framework will be generalized so that it can be used to test for differences in effective connectivity between different experimental conditions; and various populations (e.g., healthy controls vs. patients). Finally, we shall adopt the mixed effects vector autoregressive (ME-VAR) model by Gorrostieta et al. (2012) to account for the variation in effective connectivity among different subjects even in a homogeneous group. Finally, as noted, PDC appears to have evolved during resting-state which could indicate non-stationarity of the cortical sources even during resting-state. One approach will be to adapt the novel model developed in Fiecas and Ombao (2016); Samdin et al. (2016) for evolving brain processes to the high-dimensional setting.

Bibliography

Astolfi, L., Cincotti, F., Babiloni, C., Carducci, F., Basilisco, A., Rossini, P. M., Salinari, S., Mattia, D., Cerutti, S., Dayan, D. B., et al. (2005). Estimation of the cortical connectivity by high-resolution EEG and structural equation modeling: simulations and application to finger tapping data. *Biomedical Engineering, IEEE Transactions on*, 52(5):757–768.

Astolfi, L., Cincotti, F., Mattia, D., de Vico Fallani, F., Tocci, A., Colosimo, A., Salinari, S., Marciani, M. G., Hesse, W., Witte, H., et al. (2008). Tracking the time-varying cortical connectivity patterns by adaptive multivariate estimators. *Biomedical Engineering, IEEE Transactions on*, 55(3):902–913.

Astolfi, L., Cincotti, F., Mattia, D., Marciani, M. G., Baccala, L. A., de Vico Fallani, F., Salinari, S., Ursino, M., Zavaglia, M., Ding, L., et al. (2007a). Comparison of different cortical connectivity estimators for high-resolution EEG recordings. *Human brain mapping*, 28(2):143–157.

Astolfi, L., De Vico Fallani, F., Cincotti, F., Mattia, D., Marciani, M., Bufalari, S., Salinari, S., Colosimo, A., Ding, L., Edgar, J., et al. (2007b). Imaging functional brain connectivity patterns from high-resolution EEG and fMRI via graph theory. *Psychophysiology*, 44(6):880–893.

Babiloni, F., Cincotti, F., Babiloni, C., Carducci, F., Mattia, D., Astolfi, L., Basilisco, A., Rossini, P., Ding, L., Ni, Y., et al. (2005). Estimation of the cortical functional connectivity with the multimodal integration of high-resolution EEG and fMRI data by directed transfer function. *NeuroImage*, 24(1):118–131.

Baccalá, L. A. and Sameshima, K. (2001a). Partial directed coherence: a new concept in neural structure determination. *Biological cybernetics*, 84(6):463–474.

Baccalá, L. A. and Sameshima, K. (2001b). Partial directed coherence: a new concept in neural structure determination. *Biological cybernetics*, 84(6):463–474.

Bai, J. and Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, 70(1):191–221.

Baillet, S., Mosher, J. C., and Leahy, R. M. (2001a). Electromagnetic brain mapping. *Signal Processing Magazine, IEEE*, 18(6):14–30.

Baillet, S., Mosher, J. C., and Leahy, R. M. (2001b). Electromagnetic brain mapping. *Signal Processing Magazine, IEEE*, 18(6):14–30.

Basu, S., Michailidis, G., et al. (2015). Regularized estimation in sparse high-dimensional time series models. *The Annals of Statistics*, 43(4):1535–1567.

Belouchrani, A., Abed-Meraim, K., Cardoso, J.-F., and Moulines, E. (1997). A blind source separation technique using second-order statistics. *Signal Processing, IEEE Transactions on*, 45(2):434–444.

Berger, H. (1929). Über das elektrenkephalogramm des menschen. *European Archives of Psychiatry and Clinical Neuroscience*, 87(1):527–570.

Box, G. E. and Tiao, G. C. (1977). A canonical analysis of multiple time series. *Biometrika*, 64(2):355–365.

Brillinger, D. (1964). A frequency approach to the techniques of principal components, factor analysis and canonical variates in the case of stationary time series. In *Invited Paper, Royal Statistical Society Conference, Cardiff Wales.(Available at http://stat-www. berkeley. edu/users/brill/papers. html)*.

Brillinger, D. R. (2001). *Time series: data analysis and theory*, volume 36. Siam.

Britz, J., Van De Ville, D., and Michel, C. M. (2010). BOLD correlates of EEG topography reveal rapid resting-state network dynamics. *NeuroImage*, 52(4):1162–1170.

Calhoun, V. D. and Adali, T. (2012). Multisubject independent component analysis of fMRI: a decade of intrinsic networks, default mode, and neurodiagnostic discovery. *IEEE reviews in biomedical engineering*, 5:60–73.

Calhoun, V. D., Miller, R., Pearlson, G., and Adalı, T. (2014). The chronnectome: time-varying connectivity networks as the next frontier in fMRI data discovery. *Neuron*, 84(2):262–274.

Cardoso, J.-F. (1999). High-order contrasts for independent component analysis. *Neural computation*, 11(1):157–192.

Chen, A. C., Feng, W., Zhao, H., Yin, Y., and Wang, P. (2008). EEG default mode network in the human brain: spectral regional field powers. *NeuroImage*, 41(2):561–574.

Cheung, B. L. P., Riedner, B. A., Tononi, G., and van Veen, B. (2010). Estimation of cortical connectivity from EEG using state-space models. *Biomedical Engineering, IEEE Transactions on*, 57(9):2122–2134.

Chiang, J., Wang, Z. J., and McKeown, M. J. (2009). Sparse multivariate autoregressive (MAR)-based partial directed coherence (PDC) for electroencephalogram (EEG) analysis. In *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on*, pages 457–460. IEEE.

Chiang, J., Wang, Z. J., and McKeown, M. J. (2012). A generalized multivariate autoregressive (GmAR)-based approach for EEG source connectivity analysis. *Signal Processing, IEEE Transactions on*, 60(1):453–465.

Cribben, I., Haraldsdottir, R., Atlas, L. Y., Wager, T. D., and Lindquist, M. A. (2012). Dynamic connectivity regression: determining state-related changes in brain connectivity. *NeuroImage*, 61(4):907–920.

Davis, R., Zang, P., and Zheng, T. (0). Sparse vector autoregressive modeling. *Journal of Computational and Graphical Statistics*, 0(ja):1–53.

Davis, R. A., Zang, P., and Zheng, T. (2015). Sparse vector autoregressive modeling. *Journal of Computational and Graphical Statistics*, (just-accepted):1–53.

de Peralta Menendez, R. G., Andino, S. G., Lantz, G., Michel, C. M., and Landis, T. (2001). Noninvasive localization of electromagnetic epileptic activity. I. Method descriptions and simulations. *Brain topography*, 14(2):131–137.

Ding, L., Wilke, C., Xu, B., Xu, X., van Drongelene, W., Kohrman, M., and He, B. (2007). EEG source imaging: correlate source locations and extents with ECoG and surgical resections in epilepsy patients. *Journal of clinical neurophysiology: official publication of the American Electroencephalographic Society*, 24(2):130.

Durbin, J. and Koopman, S. (2001). Time Series Analysis by Space State Methods. *Nova Iorque: Oxford University Press*.

et al., J. V. (2008). Evidence for a frontoparietal control system revealed by intrinsic functional connectivity. *J. Neurophysiol.*, 100:3328–3342.

et al., M. E. R. (2001). A default mode of brain function. *Proc. Natl. Acad. Sci. U. S. A.*, 98:676–682.

et al., R. L. (2011). Large-scale directional connections among multi resting-state neural networks in human brain: A functional MRI and Bayesian network modeling study. *NeuroImage*, 56:1035–1042.

Escoufier, Y. (1973). Le traitement des variables vectorielles. *Biometrics*, pages 751–760.

Fiecas, M. and Ombao, H. (2011). The generalized shrinkage estimator for the analysis of functional connectivity of brain signals. *The Annals of Applied Statistics*, pages 1102–1125.

Fiecas, M. and Ombao, H. (2016). Modeling the evolution of dynamic brain processes during an associative learning experiment. *Journal of the American Statistical Association*, In press.

Fiecas, M., Ombao, H., van Lunen, D., Baumgartner, R., Coimbra, A., and Feng, D. (2013). Quantifying temporal correlations: A test–retest evaluation of functional connectivity in resting-state fMRI. *NeuroImage*, 65:231–241.

Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Ségonne, F., Salat, D. H., Busa, E., Seidman, L. J., Goldstein, J., Kennedy, D., et al. (2004). Automatically parcellating the human cerebral cortex. *Cerebral cortex*, 14(1):11–22.

Forni, M., Hallin, M., Lippi, M., and Reichlin, L. (2012). The generalized dynamic factor model. *Journal of the American Statistical Association*.

Fransson, P. and Marrelec, G. (2008). The precuneus/posterior cingulate cortex plays a pivotal role in the default mode network: Evidence from a partial correlation network analysis. *NeuroImage*, 42(3):1178–1184.

Friston, K. J. (1994). Functional and effective connectivity in neuroimaging: a synthesis. *Human brain mapping*, 2(1-2):56–78.

Friston, K. J. et al. (1994). Functional and effective connectivity in neuroimaging: a synthesis. *Human brain mapping*, 2(1-2):56–78.

Friston, K. J., Harrison, L., and Penny, W. (2003). Dynamic causal modelling. *NeuroImage*, 19(4):1273–1302.

Geweke, J. (1982). Measurement of linear dependence and feedback between multiple time series. *Journal of the American statistical association*, 77(378):304–313.

Gómez-Herrero, G., Atienza, M., Egiazarian, K., and Cantero, J. L. (2008). Measuring directional coupling between EEG sources. *NeuroImage*, 43(3):497–508.

Gorodnitsky, I. F., George, J. S., and Rao, B. D. (1995). Neuromagnetic source imaging with FOCUSS: a recursive weighted minimum norm algorithm. *Electroencephalography and clinical Neurophysiology*, 95(4):231–251.

Gorrostieta, C., Ombao, H., Bédard, P., and Sanes, J. N. (2012). Investigating brain connectivity using mixed effects vector autoregressive models. *NeuroImage*, 59(4):3347–3355.

Gott, A. N., Eckley, I. A., and Aston, J. A. (2015). Estimating the population local wavelet spectrum with application to non-stationary functional magnetic resonance imaging time series. *Statistics in medicine*, 34(29):3901–3915.

Gramfort, A., Papadopoulo, T., Olivi, E., Clerc, M., et al. (2010). OpenMEEG: opensource software for quasistatic bioelectromagnetics. *Biomed. Eng. Online*, 9(1):45.

Granger, C. W. (1969a). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, pages 424–438.

Granger, C. W. (1969b). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, pages 424–438.

Granger, C. W. (1980). Testing for causality: a personal viewpoint. *Journal of Economic Dynamics and control*, 2:329–352.

Granger, C. W. (1988). Some recent development in a concept of causality. *Journal of econometrics*, 39(1):199–211.

Grech, R., Cassar, T., Muscat, J., Camilleri, K. P., Fabri, S. G., Zervakis, M., Xanthopoulos, P., Sakkalis, V., and Vanrumste, B. (2008). Review on solving the inverse problem in EEG source analysis. *Journal of neuroengineering and rehabilitation*, 5(1):25.

Grill-Spector, K., Henson, R., and Martin, A. (2006). Repetition and the brain: neural models of stimulus-specific effects. *Trends in cognitive sciences*, 10(1):14–23.

Haas, L. (2003). Hans Berger (1873–1941), Richard Caton (1842–1926), and electroencephalography. *Journal of Neurology, Neurosurgery & Psychiatry*, 74(1):9–9.

Hämäläinen, M. S. and Ilmoniemi, R. (1994a). Interpreting magnetic fields of the brain: minimum norm estimates. *Medical & biological engineering & computing*, 32(1):35–42.

Hämäläinen, M. S. and Ilmoniemi, R. J. (1994b). Interpreting magnetic fields of the brain: minimum norm estimates. *Medical & biological engineering & computing*, 32(1):35–42.

Hamalainen, M. S. and Sarvas, J. (1989). Realistic conductivity geometry model of the human head for interpretation of neuromagnetic data. *Biomedical Engineering, IEEE Transactions on*, 36(2):165–171.

Harvey, A. C. (1990). *Forecasting, structural time series models and the Kalman filter*. Cambridge university press.

Hassan, M., Dufor, O., Merlet, I., Berrou, C., and Wendling, F. (2014). EEG source connectivity analysis: from dense array recordings to brain networks.

Haufe, S., Dähne, S., and Nikulin, V. V. (2014). Dimensionality reduction for the analysis of brain oscillations. *NeuroImage*, 101:583–597.

Haufe, S., Nikulin, V. V., Müller, K.-R., and Nolte, G. (2013). A critical assessment of connectivity measures for EEG data: a simulation study. *NeuroImage*, 64:120–133.

Haufe, S., Tomioka, R., Nolte, G., M´uller, K.-R., and Kawanabe, M. (2010). Modeling sparse connectivity between underlying brain sources for EEG/MEG. *Biomedical Engineering, IEEE Transactions on*, 57(8):1954–1963.

He, B., Yang, L., Wilke, C., and Yuan, H. (2011). Electrophysiological imaging of brain activity and connectivity-challenges and opportunities. *IEEE Trans. Biomed. Eng.*, 58(7):1918–1931.

Holmes, C. J., Hoge, R., Collins, L., Woods, R., Toga, A. W., and Evans, A. C. (1998). Enhancement of MR images using registration for signal averaging. *Journal of computer assisted tomography*, 22(2):324–333.

Hui, H. B., Pantazis, D., Bressler, S. L., and Leahy, R. M. (2010). Identifying true cortical interactions in MEG using the nulling beamformer. *NeuroImage*, 49(4):3161–3174.

Jung, T.-P., Makeig, S., Humphries, C., Lee, T.-W., Mckeown, M. J., Iragui, V., and Sejnowski, T. J. (2000). Removing electroencephalographic artifacts by blind source separation. *Psychophysiology*, 37(02):163–178.

Kaminski, M. and Blinowska, K. J. (1991). A new method of the description of the information flow in the brain structures. *Biological cybernetics*, 65(3):203–210.

Kang, H., Ombao, H., Linkletter, C., Long, N., and Badre, D. (2012). Spatio-spectral mixed-effects model for functional magnetic resonance imaging data. *Journal of the American Statistical Association*, 107(498):568–577.

Khanna, A., Pascual-Leone, A., Michel, C. M., and Farzan, F. (2015). Microstates in resting-state EEG: Current status and future directions. *Neuroscience & Biobehavioral Reviews*, 49:105–113.

Knyazev, G. G., Slobodskoj-Plusnin, J. Y., Bocharov, A. V., and Pylkova, L. V. (2011). The default mode network and EEG alpha oscillations: an independent component analysis. *Brain research*, 1402:67–79.

Korobilis, D. (2013). VAR forecasting using Bayesian variable selection. *Journal of Applied Econometrics*, 28(2):204–230.

Kybic, J., Clerc, M., Abboud, T., Faugeras, O., Keriven, R., and Papadopoulo, T. (2005). A common formalism for the integral formulations of the forward EEG problem. *Medical Imaging, IEEE Transactions on*, 24(1):12–28.

Lam, C. and Yao, Q. (2012). Factor modeling for high-dimensional time series: inference for the number of factors. *The Annals of Statistics*, pages 694–726.

Lazar, N. (2008). *The statistical analysis of functional MRI data*. Springer Science & Business Media.

Lila, E., Aston, J. A., and Sangalli, L. M. (2016). Smooth Principal Component Analysis over two-dimensional manifolds with an application to Neuroimaging. *arXiv preprint arXiv:1601.03670*.

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

Luiz A. Baccala, K. S. (2014). Partial Directed Coherence. In Koichi Sameshima, L. A. B., editor, *Methods in Brain Connectivity Inference through Multivariate Time Series Analysis*, chapter 4, pages 57–72. CRC Press.

Makeig, S., Bell, A. J., Jung, T.-P., Sejnowski, T. J., et al. (1996a). Independent component analysis of electroencephalographic data. *Advances in neural information processing systems*, pages 145–151.

Makeig, S., Bell, A. J., Jung, T.-P., Sejnowski, T. J., et al. (1996b). Independent component analysis of electroencephalographic data. *Advances in neural information processing systems*, pages 145–151.

Mantini, D., Perrucci, M. G., Del Gratta, C., Romani, G. L., and Corbetta, M. (2007). Electrophysiological signatures of resting state networks in the human brain. *Proc. Natl. Acad. Sci. U. S. A.*, 104(32):13170–13175.

Matsuura, K. and Okabe, Y. (1995). Selective minimum-norm solution of the biomagnetic inverse problem. *Biomedical Engineering, IEEE Transactions on*, 42(6):608–615.

Meunier, D., Achard, S., Morcom, A., and Bullmore, E. (2009). Age-related changes in modular organization of human brain functional networks. *NeuroImage*, 44(3):715–723.

Meunier, D., Lambiotte, R., and Bullmore, E. T. (2010). Modular and hierarchically modular organization of brain networks. *Frontiers in neuroscience*, 4.

Michel, C. M., Murray, M. M., Lantz, G., Gonzalez, S., Spinelli, L., and de Peralta, R. G. (2004). EEG source imaging. *Clinical neurophysiology*, 115(10):2195–2222.

Mørup, M., Hansen, L. K., Arnfred, S. M., Lim, L.-H., and Madsen, K. H. (2008). Shift-invariant multilinear decomposition of neuroimaging data. *NeuroImage*, 42(4):1439–1450.

Nedungadi, A. G., Ding, M., and Rangarajan, G. (2011). Block coherence: a method for measuring the interdependence between two blocks of neurobiological time series. *Biological cybernetics*, 104(3):197–207.

Ngo, D., Sun, Y., Genton, M. G., Wu, J., Srinivasan, R., Cramer, S. C., and Ombao, H. (2015). An exploratory data analysis of electroencephalograms using the functional boxplots approach. *Frontiers in neuroscience*, 9.

Nicholson, W., Matteson, D., and Bien, J. (2014). Structured regularization for large vector autoregressions. *Cornell University*.

Niedermeyer, E. and da Silva, F. L. (2005). *Electroencephalography: basic principles, clinical applications, and related fields*. Lippincott Williams & Wilkins.

Nolte, G., Ziehe, A., Nikulin, V. V., Schlögl, A., Krämer, N., Brismar, T., and Müller, K.-R. (2008). Robustly estimating the flow direction of information in complex physical systems. *Physical review letters*, 100(23):234101.

Nunez, P. L. and Srinivasan, R. (2006a). *Electric fields of the brain: the neurophysics of EEG*. Oxford university press.

Nunez, P. L. and Srinivasan, R. (2006b). *Electric fields of the brain: the neurophysics of EEG*. Oxford university press.

Nunez, P. L., Srinivasan, R., Westdorp, A. F., Wijesinghe, R. S., Tucker, D. M., Silberstein, R. B., and Cadusch, P. J. (1997). EEG coherency: I: statistics, reference electrode, volume conduction, Laplacians, cortical imaging, and interpretation at multiple scales. *Electroencephalography and clinical neurophysiology*, 103(5):499–515.

Ombao, H. and Van Bellegem, S. (2008). Evolutionary coherence of nonstationary signals. *IEEE Transactions on Signal Processing*, 56(6):2259–2266.

Omidvarnia, A., Azemi, G., Boashash, B., O'Toole, J. M., Colditz, P. B., and Vanhatalo, S. (2014). Measuring time-varying information flow in scalp eeg signals: orthogonalized partial directed coherence. *Biomedical Engineering, IEEE Transactions on*, 61(3):680–693.

Onton, J., Westerfield, M., Townsend, J., and Makeig, S. (2006). Imaging human EEG dynamics using independent component analysis. *Neuroscience & Biobehavioral Reviews*, 30(6):808–822.

Pascual-Marqui, R., Biscay, R., Bosch-Bayard, J., Lehmann, D., Kochi, K., Yamada, N., Kinoshita, T., and Sadato, N. (2014). Isolated effective coherence (iCoh): causal information flow excluding indirect paths. *arXiv preprint arXiv:1402.4887*.

Pascual-Marqui, R. D. (2007a). Coherence and phase synchronization: generalization to pairs of multivariate time series, and removal of zero-lag contributions. *arXiv preprint arXiv:0706.1776*.

Pascual-Marqui, R. D. (2007b). Instantaneous and lagged measurements of linear and nonlinear dependence between groups of multivariate time series: frequency decomposition. *arXiv preprint arXiv:0711.1455*.

Pascual-Marqui, R. D. et al. (2002). Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details. *Methods Find Exp Clin Pharmacol*, 24(Suppl D):5–12.

Peña, D. and Yohai, V. J. (2015). Generalized Dynamic Principal Components. *Journal of the American Statistical Association*, (just-accepted):00–00.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1985). Learning internal representations by error propagation. Technical report, DTIC Document.

Rush, S. and Driscoll, D. A. (1969). EEG electrode sensitivity-an application of reciprocity. *Biomedical Engineering, IEEE Transactions on*, (1):15–22.

Samdin, S. B., Ting, C.-M., Ombao, H., and Salleh, S.-H. (2016). A unified estimation framework for state-related changes in effective brain connectivity. *IEEE Trans. Biomedical Engineering*.

Samworth, R. J., Yuan, M., et al. (2012). Independent component analysis via nonparametric maximum likelihood estimation. *The Annals of Statistics*, 40(6):2973–3002.

Sato, J. R., Fujita, A., Cardoso, E. F., Thomaz, C. E., Brammer, M. J., and Amaro, E. (2010). Analyzing the connectivity between regions of interest: an approach based on cluster Granger causality for fMRI data analysis. *NeuroImage*, 52(4):1444–1455.

Schreiber, T. (2000). Measuring information transfer. *Physical review letters*, 85(2):461.

Schwartz, R. S., Brown, E. N., Lydic, R., and Schiff, N. D. (2010). General anesthesia, sleep, and coma. *New England Journal of Medicine*, 363(27):2638–2650.

Shumway, R. H. and Stoffer, D. S. (2010). *Time series analysis and its applications: with R examples*. Springer Science & Business Media.

Smith, S. M., Beckmann, C. F., Andersson, J., Auerbach, E. J., Bijsterbosch, J., Douaud, G., Duff, E., Feinberg, D. A., Griffanti, L., Harms, M. P., et al. (2013). Resting-state fMRI in the human connectome project. *NeuroImage*, 80:144–168.

Srinivasan, R. and Deng, S. (2012). *Multivariate Spectral Analysis of Electroencephalogram*, pages 1–24. CRC Press.

Stock, J. H. and Watson, M. W. (2002). Forecasting using principal components from a large number of predictors. *Journal of the American statistical association*, 97(460):1167–1179.

Tadel, F., Baillet, S., Mosher, J. C., Pantazis, D., and Leahy, R. M. (2011). Brainstorm: a user-friendly application for MEG/EEG analysis. *Computational intelligence and neuroscience*, 2011:8.

Teplan, M. (2002). Fundamentals of EEG measurement. *Measurement science review*, 2(2):1–11.

Ting, C.-M., Seghouane, A.-K., Salleh, S.-H., and Noor, A. M. (2014). Estimating effective connectivity from fmri data using factor-based subspace autoregressive models. *IEEE Sig. Process. Lett.*, 22(6):757–761.

Uutela, K., Hamalainen, M., and Salmelin, R. (1998). Global optimization in the localization of neuromagnetic sources. *Biomedical Engineering, IEEE Transactions on*, 45(6):716–723.

Uutela, K., Hämäläinen, M., and Somersalo, E. (1999). Visualization of magnetoencephalographic data using minimum current estimates. *NeuroImage*, 10(2):173–180.

van den Heuvel, M. P. and Hulshoff, P. H. E. (2010). Exploring the brain network: A review on resting-state fMRI functional connectivity. *European Neuropsychopharmacol.*, 20:519–534.

Wang, Y., Hu, L., and Ombao, H. (2016a). *Handbook of NeuroImaging Data Analysis*, chapter Statistical Analysis of Electroencephagrams. CRC Press, New York.

Wang, Y., Ting, C.-M., and Ombao, H. (2016b). Modeling Effective Connectivity in High-Dimensional Cortical Source Signals. *IEEE Journal of Selected Topics in Signal Processing*, 10(7):1315.

Wu, J., Srinivasan, R., Kaur, A., and Cramer, S. C. (2014). Resting-state cortical connectivity predicts motor skill acquisition. *NeuroImage*, 91:84–90.

Yan, Y., Nunez, P., and Hart, R. (1991). Finite-element model of the human head: scalp potentials due to dipole sources. *Medical and Biological Engineering and Computing*, 29(5):475–481.

Zhou, Z., Ding, M., Chen, Y., Wright, P., Lu, Z., and Liu, Y. (2009). Detecting directional influence in fMRI connectivity analysis using PCA based Granger causality. *Brain research*, 1289:22–29.

Appendix A

Appendix

A.1 Algorithms

Algorithm 1 Compute Spectral Matrix

procedure CrossPowerSpectrum($\{\mathbf{z}(t)\}$)
  set nfft $= T$, $m = [\sqrt{T}]$, $h_\ell = \frac{1}{2m+1}$
  //Remark 1: $h_\ell \geq 0$, $h_{-\ell} = h_\ell$ and $\sum_\ell h_\ell = 1$
  //Remark 2: In order to make $\widehat{S}_{zz}$ positive definite we need $2m + 1 > p$
  for $k = 0, \ldots, T-1$ **do**
    compute $\mathbf{z}_\omega(k) \leftarrow \sum_{t=0}^{T-1} \mathbf{z}(t) \exp(-i2\pi t \frac{k}{T})$ ▷ transfer function: fft
    compute $I_\omega(k) \leftarrow \mathbf{z}_\omega(k)\mathbf{z}_\omega^*(k)$ ▷ raw periodogram
  end for
  for $k = 0, \ldots, m$ **do** ▷ padding
    $I_\omega(-k) \leftarrow I_\omega^*(k)$
    $I_\omega(T+k) \leftarrow I_\omega^*(T-k)$
  end for
  for $k = 0, \ldots, T-1$ **do** ▷ smoothing
    compute $\widehat{S}_{zz}(k) \leftarrow \sum_{\ell=-m}^{m} h_\ell I_\omega(k+\ell)$
  end for
  return $\{\mathbf{z}_\omega(k)\}$, $\{\widehat{S}_{zz}(k)\}$
end procedure

Algorithm 2 Factor model - m components

procedure PCA($\{\widehat{S}_{zz}(k)\}$)
  for $k = 0, \ldots, T-1$ **do**
    compute eigenvalues $\lambda_1(k) > \lambda_2(k), \ldots, \lambda_n(k)$
    compute corresponding eigenvectors $\mathbf{e}_1(k), \mathbf{e}_2(k), \ldots, \mathbf{e}_n(k)$
  end for
  return $\{\lambda_j(k)\}$, $\{\mathbf{e}_j(k)\}$
end procedure
procedure ComputeFactors($\{\mathbf{z}_\omega(k)\}$, $\{\lambda_j(k)\}$, $\{\mathbf{e}_j(k)\}$)
  for $k = 0, \ldots, \lfloor T/2 \rfloor$ **do**
    $A(k) \leftarrow [\mathbf{e}_1(k), \ldots, \mathbf{e}_m(k)]$
  end for
  for $k = \lfloor T/2 \rfloor + 1, \ldots, T-1$ **do**
    $A(k) \leftarrow A^*(T-k)$
  end for
  for $k = 0, \ldots, T-1$ **do**
    $\mathbf{f}_\omega(k) = A^*(T-k)\mathbf{z}_\omega(k)$ ▷ transfer function for $\mathbf{f}(t)$
  end for
  for $t = 0, \ldots, T-1$ **do**
    $\mathbf{f}(t) = \frac{1}{T} \sum_{k=0}^{T-1} \mathbf{f}_\omega(k) \exp(i2\pi t \frac{k}{T})$ ▷ ifft
  end for
  return $\{\mathbf{f}(t)\}$, $\{A(k)\}$
end procedure