Modeling Effective Connectivity in High-Dimensional Cortical Source Signals

Yuxiao Wang, Chee-Ming Ting, Member, IEEE and Hernando Ombao*

Abstract—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.

Index Terms—Coherence analysis; Dimension reduction; Factor analysis; Multichannel EEG; Partial directed coherence; Principal component analysis; Vector autoregressive model.

I. 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 [1]. Effective and functional connectivity are both usually inferred from brain signals such as functional magnetic resonance imaging (fMRI) and multichannel electroencephalogram (EEG) [2]. 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 [3]–[6]. However, the EEG signal captured by a specific sensor on the scalp is not a direct measurement of 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 [17], 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 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 averaging signals [17], our model is more comprehensive because it is able to characterize connectivity at both dipole level and region level.
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) \([20]\) and the VAR+ICA model formulated in the state-space form \([21]–[24]\). 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 \([25]\). For a meshgrid of cortex area with 3 mm spacing, the number of grid points (equivalent dipoles) can be greater than \(10^3\), 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 \([26]\) 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 \([9], [27]\). 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.

In this paper, 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 \([28]\) 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) \([29]\) 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 \([30]\). 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 \([31]\).

To establish directed dependence between ROIs, the temporal dependency structure of combined 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 paper 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.

II. 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.

A. EEG Signal Model

According to biophysical models, scalp EEG signals can be modeled as a linear mixing of sources on the cortex. Let \(\mathbf{Z}(t) = [\mathbf{Z}_1(t), \ldots, \mathbf{Z}_N(t)]' \in \mathbb{R}^N, t = 1, \ldots, T\) 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

\[
\mathbf{X}(t) = \mathbf{A}\mathbf{Z}(t) + \mathbf{N}(t) \tag{1}
\]
where \( N(t) \) is the measurement and physiological noise; \( A \in \mathbb{R}^{M \times N} \) is the mixing matrix (which is also known as the lead field matrix). The above model for EEG has been widely used in EEG source reconstruction and studies on EEG source connectivity [32]-[34].

B. Cortical Source Models

1) Factor Model for a Cortical Region: The entire cortex area can be parcelated 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 \( Z_r(t) \in \mathbb{R}^{n_r} \) 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 \( Z_r(t) \) at region \( r \) to be driven by factor activity \( f_r(t) \in \mathbb{R}^{m_r} \) with loading matrix \( Q_r \), where the dimension of \( f_r(t) \) is much lower than that of \( Z_r(t) \), i.e. \( m_r \ll n_r \). The factor model at region \( r \) is given by

\[
Z_r(t) = Q_r f_r(t)
\]

(2)

where \( Q_r \in \mathbb{R}^{n_r \times m_r} \) is the factor loading matrix for region \( r \) and \( f_r(t) \) represents factor activity at time \( t \). For identifiability or unique model decomposition, we assume the following constraints on both \( Q_r \) and \( f_r(t) \): \( Q_r' Q_r = I_{m_r} \) and \( \text{Cov}(f_r(t)) \) is a diagonal matrix with distinct positive diagonal elements.

2) Factor Model for the Entire Cortex: The source activity of the whole cortex area denoted by \( Z(t) = [Z_1(t)' , Z_2(t)' , \ldots , Z_R(t)']' \) can be represented by a global factor model

\[
Z(t) = \sum_{\ell=1}^{P} \Phi^{(\ell)}(t) f(t-\ell) + \eta(t)
\]

(3)

where \( f(t) = [f_1(t) , f_2(t) , \ldots , f_R(t)]' \in \mathbb{R}^M \) and \( \Phi = \text{diag}\{Q_1 , Q_2 , \ldots , Q_R\} \) 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.

3) VAR Model for Factor Activity: According to the factor model (3), \( Z(t) \) is the instantaneous mixing of the factor activity \( f(t) \). Therefore the temporal inter-dependence structure for \( Z(t) \) can be characterized by the temporal dependence structure of factor \( f(t) \). We use a VAR model with order \( P \) (denoted \( \text{VAR}(P) \)) to characterize the temporal dependence structure in \( f(t) \)

\[
f(t) = \sum_{\ell=1}^{P} \Phi^{(\ell)} f(t-\ell) + \eta(t)
\]

(4)

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

\[
\Phi^{(\ell)} = \begin{pmatrix}
\Phi_{f_1f_1}(\ell) & \ldots & \Phi_{f_1f_R}(\ell) \\
\vdots & \ddots & \vdots \\
\Phi_{f_Rf_1}(\ell) & \ldots & \Phi_{f_Rf_R}(\ell)
\end{pmatrix}
\]

The diagonal block \( \Phi_{f_kf_k}(\ell) \) summarizes the lagged autocorrelation within each region, while the off-diagonal block \( \Phi_{f_kf_l}(\ell) \), where \( j \neq k \), captures the cross-dependence between regions \( j \) and \( k \); and \( \eta(t) \in \mathbb{R}^M \) represents the Gaussian white noise with mean zeros and covariance matrix \( \Sigma_n \).

4) 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 \( N \times N \) 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 \( Z(t) \), 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 [33]. We use a VAR model to characterize the temporal inter-dependence structure for the factor activity \( f(t) \), as described in Section II-B3. We show that the temporal interdependence structure for the high-dimensional source activity \( Z(t) \) 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 \( Z(t) = Q f(t) \) (i.e., the significant activity in the sources are captured by the factors). Substituting the VAR factor process (4) in the model defined in (3) gives

\[
Z(t) = \sum_{\ell=1}^{P} \Phi^{(\ell)} f(t-\ell) + \eta(t)
\]

(5)

\[
= \sum_{\ell=1}^{P} Q \Phi^{(\ell)} Q' f(t-\ell) + Q \eta(t)
\]

(6)

\[
= \sum_{\ell=1}^{P} Q \Phi^{(\ell)} Q' Z(t-\ell) + Q \eta(t).
\]

(7)

Finally, we have the following factor VAR model for \( Z(t) \)

\[
Z(t) = \sum_{\ell=1}^{P} \Phi^{(\ell)} Z(t-\ell) + E(t)
\]

(8)

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

C. 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.
1) 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 $\omega$) 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 $\Phi^Z(\omega) = \mathbf{I} - \sum_{t=1}^{P} \Phi^Z(t) \exp(-i2\pi \omega t/\Omega_s)$ be the Fourier transform of the VAR coefficient matrix $\Phi^Z = [\Phi_{ij}(\omega)]_{1\leq i,j\leq N} \in \mathbb{R}^{N \times N}$ at frequency $\omega$, where $\mathbf{I}$ is the identity matrix, and $\Omega_s$ is the sampling frequency. The PDC derived from the VAR model (8) is defined as

$$\pi_{ij}(\omega) = \frac{|\Phi^Z_{ij}(\omega)|}{\sum_{k=1}^{N} \Phi^Z_{ik}(\omega)\Phi^Z_{kj}(\omega))^*}.$$  (9)

Here, $\pi_{ij}(\omega) \in [0,1]$. It gives an indication of the strength of the linear directed influence of $\omega$-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 $\pi(\omega) = [\pi_{ij}(\omega)]_{1\leq i,j\leq N} \in \mathbb{R}^{N \times N}$ characterizes a network of directed interactions between dipoles of the entire cortex at frequency $\omega$.

2) 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 $\omega$ is $S^f(\omega) = \mathbf{Q}S^\star(\omega)\mathbf{Q}^\dagger$, where $\mathbf{S}^\star(\omega)$ is the spectral matrix of $\mathbf{f}(t)$, which can be computed as

$$S^f(\omega) = \mathbf{H}^f(\omega)\Sigma_\eta(\mathbf{H}^f(\omega))^\dagger$$  (10)

where $\mathbf{H}^f(\omega) = (\Phi^f(\omega))^{-1}$. The coherence between dipoles $i$ and $j$ is then defined as

$$\rho_{ij}^2(\omega) = \frac{|S_{ij}^f(\omega)|^2}{S_{ii}^f(\omega)S_{jj}^f(\omega)}.$$  (11)

3) 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 $C_{r_1,r_2}(\omega)$ is computed as

$$C_{r_1,r_2}(\omega) = \frac{1}{|U_{r_1}| |U_{r_2}|} \sum_{i \in U_{r_1}, j \in U_{r_2}} \pi_{ij}(\omega).$$  (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 [35] or frequency decomposition of Granger causality [36].

III. 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.

A. 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 $10^3$) is smaller than the dimension of the cortex source space (which can be of order $10^4$).

1) Obtaining the Source Mixing Matrix: The lead field matrix $\mathbf{h}$ in the EEG observation model (1) can be computed by applying the boundary element method (BEM) on a discretized realistic head MRI template, as shown in Fig. 1. In this paper, we use the Colin 27 MRI head template [37]. Fig. 1. BEM head model computed from the Colin 27 head template. The BEM head model is computed using the software package Open-MEEG(compute) [38], [39], and the discretized head is visualized using the software Brainstorm [40].

2) Inverse Source Reconstruction from EEG: Given the lead field matrix $\mathbf{h}$ 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 [29]. 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) [41] method that minimizes the objective function of (13) with respect to $\mathbf{Z}(t)$ and $c$, given lead field matrix $\mathbf{h}$, EEG observation $\mathbf{X}$ and $\alpha$.

$$F = \sum_{t=1}^{T} \|\mathbf{X}(t) - \mathbf{h}\mathbf{Z}(t) - c\mathbf{1}\|^2 + \alpha\|\mathbf{Z}(t)\|^2.$$  (13)
The solution of the minimization gives the estimate \( \hat{Z}(t) = GX(t) \), where \( G = \hat{A}'H(H\hat{A}\hat{A}' + \alpha H)^{-1} \) with \( H = I - 11'/1'1'\). The regularization parameter \( \alpha \) is determined by cross-validation, which is the same method as described in [29]. The sLORETA estimation for the source is then obtained by standardizing \( \hat{Z}(t) \) using its variance estimate.

B. 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 \( \hat{Z}_r(t) \) to estimate the regional factor activity \( f_r(t) \) and its loading \( Q_r(t) \) in (2). Then we concatenate these regional factor estimates to form the estimate for the global factor model in (3) for the entire cortex. A VAR model for the factor space in (3) is fitted on the estimated global factor time series, and then projected to the VAR model in source space in (6) 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 \( \hat{Z}_r = [\hat{Z}_r(1), \ldots, \hat{Z}_r(T)]' \in \mathbb{R}^{T \times n_r} \) be the reconstructed cortical sources for region \( r \). Here are the estimation steps.

- Compute the sample covariance matrix of \( Z_r(t) \), denoted as \( \hat{\Sigma}_{Z_rZ_r} \):

\[
\hat{\Sigma}_{Z_rZ_r} = \hat{Z}_r' \hat{Z}_r / T = \frac{1}{T} \sum_{t=1}^{T} \hat{Z}_r(t) \hat{Z}_r(t)'
\]  

where \( \hat{Z}_r(t) \) has zero mean.

- Estimate the factor loading matrix \( Q_r \in \mathbb{R}^{n_r \times m_r} \) and the factors \( f_r(t) \) based on the eigenvalue-eigenvector decomposition of \( \hat{\Sigma}_{Z_rZ_r} \). Let \( \lambda_1, \ldots, \lambda_{n_r} \) be the unique eigenvalues of \( \hat{\Sigma}_{Z_rZ_r} \), in decreasing order and let \( V_1, \ldots, V_{n_r} \in \mathbb{R}^{n_r} \) be the corresponding orthonormal eigenvectors. The estimator of \( Q_r \) can be defined as

\[
\hat{Q}_r = [V_1, \ldots, V_{m_r}]
\]

where \( m_r \) is the dimension of the regional factor activity \( f_r(t) \) 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 \( f_r \in \mathbb{R}^{T \times m_r} \)

\[
\hat{f}_r = \hat{Z}_r \hat{Q}_r
\]

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

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

\[
\text{AIC}(P) = \log(\hat{\Sigma}_\eta(P)) + \frac{2}{T} PM^2
\]

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

- Compute the VAR coefficient estimate for source activity \( Z(t) \) by substitution \( \hat{\Phi}^Z(\ell) = \hat{Q}\hat{\Phi}^f(\ell)\hat{Q}' \) where \( \hat{Q} = \text{Diag}\{Q_1, \ldots, Q_{N}\} \).

- Estimate the between-dipole source connectivity by computing the PDC \( \pi_{ij}(\omega) \) from dipole \( j \) to dipole \( i \) via substitution of \( \hat{\Phi}^Z(\ell) \) in (9).

- Estimate the between-ROI source connectivity \( \hat{C}_{r_1r_2}(\omega) \) by summarizing the estimated inter-dipole PDC \( \pi_{ij}(\omega) \) over the region from \( r_2 \) to \( r_1 \) according to (12).

IV. 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.

A. 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 \( Z(t) \) from a set of \( N = 120 \) dipoles from 6 cortical regions (20 dipoles per region, regions include LFP, 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 (8) on \( Z(t) \) to obtain the VAR coefficients matrices \( \hat{\Phi}^Z \) and the covariance matrix of residuals \( \hat{\Sigma}_F \). 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 (8). The \( \hat{\Phi}^Z \) was assumed known in this simulation study and used as ground-truth to compare different estimators. We repeated the simulation for 100 times.

B. Model Evaluation

We evaluated the accuracy of the model in estimating the directed connectivity in terms of PDC in the source space. Let \( \pi(\omega) \) be the connectivity matrix whose elements are given in \( \pi_{ij}(\omega) \) (as defined in (9)) and \( \pi(\omega) \) 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 $\Lambda$, is computed as follows

$$\Lambda(\Omega) = \frac{1}{|\Omega|} \sum_{\omega \in \Omega} \frac{\|\hat{\pi}(\omega) - \pi(\omega)\|_F}{\|\pi(\omega)\|_F},$$

where $\Omega$ 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 $\|H\|_F = \text{tr}(HH')^{1/2}$ denotes the Frobenius norm of matrix $H$.

### C. 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 [42]. 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. 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.

### V. 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.
B. Resting-state Cortical Connectivity

While resting-state brain connectivity networks have been extensively investigated in fMRI studies [44], most studies on EEG focus on task-based source connectivity (e.g., motor-imagery for brain-computer interface [31]; and visual task [23]). There are few studies on the resting-state brain networks based on EEG, e.g. in [45], [46] 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 III-A1. 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.

1) 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. 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 I. 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, \( M \ll N \).

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 networks [47], [48]. 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.

![Image](312x398 to 579x519)

Fig. 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 \( \log(\hat{r}_{ij}) \), where \( i, j = 1, \ldots, N \) with \( N = 2800 \)

2) Effective Connectivity Between Regions: Fig. 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 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.

![Image](312x542 to 579x664)

<table>
<thead>
<tr>
<th>Region</th>
<th>( m_x )</th>
<th>Region</th>
<th>( m_x )</th>
<th>Region</th>
<th>( m_x )</th>
</tr>
</thead>
<tbody>
<tr>
<td>LPF</td>
<td>3-3</td>
<td>LC</td>
<td>3-3</td>
<td>RO</td>
<td>3-4</td>
</tr>
<tr>
<td>RPF</td>
<td>4-5</td>
<td>RC</td>
<td>3-3</td>
<td>LL</td>
<td>3-3</td>
</tr>
<tr>
<td>LF</td>
<td>4-5</td>
<td>LF</td>
<td>3-4</td>
<td>RL</td>
<td>4-4</td>
</tr>
<tr>
<td>RF</td>
<td>3-4</td>
<td>RP</td>
<td>4-5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LT</td>
<td>3-3</td>
<td>RT</td>
<td>2-3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table I**

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.
or information flows between the limbic, prefrontal, frontal and parietal regions (e.g., RL-LPF, RL-RPF, RL-LF and RP-RPF), which belongs to the DMN [49], 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 [50]. We also detected strong connectivity between regions of the attentional networks which include the frontal-parietal sub-network [51] (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 [52], [53].

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 [54], [55]. 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 [54] while [55] 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.

3) Time-Evolving Connectivity Across Epochs: Previous studies on the EEG source connectivity analysis assume temporal stationarity over the entire time course of recordings [20], [22], [23]. 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. 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 $R \times (j - 1) + i$. 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. 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 [56], we analyzed state-related changes in the connectivity maps directly, between underlying cortical sources.

VI. 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...
between the high-dimensional EEG sensors (rather than the sources), we will investigate generalized PDC (gPDC) [58] and generalized orthogonalized PDC (gOPDC) [59], 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 [60] 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 adopt the novel model developed in [61, 62] for evolving brain processes to the high-dimensional setting.

REFERENCES


Yuxiao Wang is a PhD candidate in the Department of Statistics at University of California at Irvine. He has been with the UCI Space-Time Modeling Group since 2012. He received the B.S. degree in physics form University of Science and Technology of China (USTC) in 2010. His current research focuses on modeling high-dimensional time series of complex structures, with applications on brain imaging analysis. He is also actively working on building software toolkits for high-dimensional time series data analysis and visualization.

Chee-Ming Ting (M’13) received the B.Eng. and M.Eng. degrees both in electrical engineering, and the Ph.D. degree in mathematics from the Universiti Teknologi Malaysia (UTM), Johor, Malaysia, in 2005, 2007 and 2012, respectively.

From 2012 to 2013, he was a Postdoctoral Fellow with the Center for Biomedical Engineering, UTM, where he is currently a Research Fellow. He was a Visiting Researcher at the University of Melbourne, Australia, and the University of California, Irvine, in 2014 and 2015, respectively. He is currently a Senior Lecturer in the Faculty of Biosciences and Medical Engineering, UTM. His research interests include statistical signal processing, spatio-temporal modeling, state-space methods, time series analysis, and high-dimensional statistics with applications to biomedical signal processing.

Dr. Ting is a member of the IEEE Signal Processing Society.

Hernando C. Ombao received the B.S. degree in mathematics from the University of the Philippines, Manila, Philippines, the M.Sc. degree in statistics from the University of California, Davis, Davis, CA, USA, and the Ph.D. degree in biostatistics from the University of Michigan, Ann Arbor, MI, USA.

He is currently a Professor in the Department of Statistics with an adjunct appointment in the Department of Cognitive Sciences, University of California at Irvine, Irvine, CA, USA, where he leads the research group on space-time modeling. He held faculty positions at the University of Pittsburgh, University of Illinois at Urbana-Champaign, and Brown University. He has served as an Associate Editor for the Journal of the American Statistical Association: Theory and Methods since 2005, Journal of the Royal Statistical Society Series B from 2011 to 2015, and the Statistical Analysis and Data Mining from 2010 to 2012. He is Co-Editor of Handbook on NeuroImaging Data Analysis, CRC Press. His work has received continuous funding from the National Science Foundation and the National Institutes of Health. His research interests include theory, methods, and models for nonstationary multivariate time series and their applications to brain signals and images.

Prof. Ombao is a Fellow of the American Statistical Association.