A Tensor Decomposition-Based Approach for Detecting Dynamic Network States From EEG

Arash Golibagh Mahyari*, David M. Zoltowski, Edward M. Bernat, and Selin Aviyente

Abstract—Functional connectivity (FC), defined as the statistical dependency between distinct brain regions, has been an important tool in understanding cognitive brain processes. Most of the current works in FC have focused on the assumption of temporally stationary networks. However, recent empirical work indicates that FC is dynamic due to cognitive functions. *Goal*: The purpose of this paper is to understand the dynamics of FC for understanding the formation and dissolution of networks of the brain. *Method*: In this paper, we introduce a two-step approach to characterize the dynamics of functional connectivity networks (FCNs) by first identifying change points at which the network connectivity across subjects shows significant changes and then summarizing the FCNs between consecutive change points. The proposed approach is based on a tensor representation of FCNs across time and subjects yielding a four-mode tensor. The change points are identified using a subspace distance measure on low-rank approximations to the tensor at each time point. The network summarization is then obtained through tensor-matrix projections across the subject and time modes. *Results*: The proposed framework is applied to electroencephalogram (EEG) data collected during a cognitive control task. The detected change-points are consistent with *a priori* known ERN interval. The results show significant connectivities in medial-frontal regions which are consistent with widely observed ERN amplitude measures. *Conclusion*: The tensor-based method outperforms conventional matrix-based methods such as singular value decomposition in terms of both change-point detection and state summarization. *Significance*: The proposed tensor-based method captures the topological structure of FCNs which provides more accurate change-point-detection and state summarization.

Index Terms—Complex networks, dynamic functional connectivity (dFC), electroencephalography, phase synchrony, tensor decomposition.

NOMENCLATURE

$C_i(t,\omega)$ Complex RID-Rihaczek distribution of the signal from the ith channel.

$\Phi_i(t,\omega)$ Time and frequency dependent phase of the signal from the ith channel.

$PLV_{(i,j)}(t,\omega)$ Phase locking value between the ith and jth channels at time t and frequency ω.

$G_{(i,j)}(t)$ Connectivity value between channels i and j at time t.

$\mathcal{G}(t)$ Three-way tensor at time t corresponding to the connectivity across nodes and subjects.

$\mathcal{X}$ Tensor.

$\mathcal{X}_{i_1,i_2,i_3}$ i_1, i_2, i_3th element of the tensor.

$\mathbf{U}^{(i)}$ ith component matrix corresponding to the left singular vectors of the tensor unfolded along the ith mode.

I. INTRODUCTION

EMPIRICAL studies from electromagnetic recordings and neuroimaging suggest that human cognition arises from transient synchronization between distant and specific neural populations [1]–[3]. The interactions between these functionally distinct neurocognitive networks, also known as functional connectivity (FC), have been studied in detail through neuronal time series recorded from different sensors, e.g., electroencephalogram (EEG) [4], or from brain voxels, e.g., functional magnetic resonance imaging (fMRI) [5], [6]. Recently, methods from complex network theory have been used to describe the topological structure and attributes of functional connectivity networks (FCNs) [7], [8], where nodes correspond to distinct brain regions and edges to pairwise associations (or functional connectivity) between them [9]. Approaches built upon the assumption of stationarity of FC have produced a wealth of literature expanding our knowledge of large-scale brain networks. However, recent studies have shown that the functional connectivity in brain networks may exhibit dynamic changes within short time scales due to task demands, learning, and anesthesia [10]–[14]. Moreover, there is growing evidence suggesting that the emergence of a unified neural process is controlled by the continuous formation and destruction of functional connectivity over multiple time scales [15], [16].

Recent work in cognitive neuroscience, for example, now suggests the presence of two broad functional brain networks that directly reflect primary systems underlying goal-directed behavior: Incentive salience and cognitive control. The incentive salience network's primary function is to scan the environment for the anticipation of reward and detection of potential danger, and centrally involves the anterior cingulate cortex, amygdala, and the major dopaminergic structures of the brain [17]–[19]. The salience network is functionally integrated with other brain structures that collectively constitute the cognitive control network, notably the lateral prefrontal cortex (LPFC),

*A. Golibagh Mahyari is with the Department of Electrical and Computer Engineering, Michigan State University, East Lansing, MI 48824, USA (e-mail: mahyari@msu.edu).

D. M. Zoltowski and Selin Aviyente are with the Department of Electrical and Computer Engineering, Michigan State University.

E. M. Bernat is with the Department of Psychology at the University of Maryland.

orbitofrontal prefrontal cortex (PFC), and inferior frontal gyrus [19]–[21]. After receiving input from the salience network, the control network becomes active to guide responses, often modifying prepotent behavior [22], [23]. This guidance entails the use of mental representations of potential future outcomes including weighting the potential positive and negative outcomes of different response options [23]. Error processing studied using the error-related negativity (ERN) is understood to involve both salience and control processing, and will be the focus of an initial application of the proposed approach.

Early dynamic network measures focused on time-varying analysis of FC graphs and characterizing network dynamics [24]–[26]. Fallani et al. [27] introduced time-varying graph analysis in the context of a simple motor task and reported temporally persistent edges. Valencia et al. showed the changes in small-worldness of the FCNs across time [28]. Dimitriadis et al. [29] developed an approach for time-varying analysis of graph theoretic indices including network efficiency and small world parameter. Similarly, Chiang et al. [30] used the hidden Markov model (HMM) to track the dynamic structure of graph theoretical network measures and to quantify the relative temporal stationarity of brain network topology. Bassett et al. [31] showed the dynamic nature of the functional connectivity network's modular structure.

In recent years, two complementary approaches to dynamic functional connectivity (dFC) tracking have been proposed. The first class of approaches are based on clustering the FCNs across time and subjects assuming that the network is in one particular state at each time point and the states fluctuate across time. For example, Allen et al. [32] proposed a data-driven approach based on k-means clustering to identify "FC-states" from resting state fMRI. Similar clustering-based approaches were adopted by Dimitriadis et al. [33], [34], Jamal et al. [35] for EEG studies and by Ou et al. [36] for resting state fMRI. Jamal et al. [35], [37]–[39] used k-means clustering on time-frequency dependent phase difference data to obtain "synchrostates" whereas Dimitriadis et al. replaced k-means by the Neural Gas algorithm to obtain a codebook that facilitates the description of "FCμstate"s [34]. Ou et al., on the other hand, employ hierarchical clustering to determine the network states from vectorized connectivity matrices and then use HMM to model the dynamics [36]. Similarly, in [40] fluctuations related to spatial components, obtained through independent vector analysis over time, are investigated through clustering and Markov modeling. In another study [41], resting state EEG FC networks were grouped into coherent states through cross-correlation analysis and evolutionary clustering of FC edge time series. Although clustering approaches provide a potentially powerful method for determining the spontaneous changes in resting state FC, they assume a distinct network state at each time point instead of a network state that is both contiguous in time and made up of multiple building blocks [42]. In the second class of approaches, Leonardi et al. [43] propose a principal component analysis (PCA)-based approach to reveal characteristic fluctuations of whole-brain FC, where the eigenconnectivities that account for the most variation in the collection of dynamic FCNs across time and subjects are denoted as network states.

The retained eigenconnectivities are treated as building blocks of dynamic FC and the subject-specific time-dependent weights that describe the contribution of each eigenconnectivity is obtained through orthogonal projection. Similarly, in [44] a PCA-based approach is used to reconstruct the principal resting-state network dynamics from EEG data. The multivariate data describing time-frequency interdependencies between electrodes is analyzed through PCA to obtain low-dimensional orthogonal subspace of resting-state network dynamics. Similarly in [45], we introduced a PCA-based approach for tracking dFC in task-based EEG based on average dFC across subjects. In [42], two approaches to network state identification, i.e., clustering and PCA, have been compared for dynamic connectivity analysis of resting-state and task-based fMRI. It was shown that resting state dFC was better described by multiple FC patterns that overlap, whereas task-based dFC was better represented by sparse separated FC patterns.

While tracking the dynamic nature of FC, the existing techniques mostly assume that the timing of the various states is known, that at each time point the brain is at a discrete state, or that the current state is a weighted sum of FC-states [42]. Although these assumptions are valid for resting-state FC studies, in the case of task-based studies the FC-states are assumed to be quasi-stationary over a time interval [41], [46]. Therefore, an important extension to current approaches would be to introduce methods that can detect the time points where significant changes to network organization occur. To address this issue, Bassett et al. [31] explored the dynamic organizational changes of graphs during learning using predefined data windows spanning multiple temporal scales. Cribben et al. [47] took a different approach and considered a data-driven technique for partitioning the time course into intervals based on the underlying FC patterns. This technique, called dynamic connectivity regression, detects temporal change points in FC and estimates a connectivity graph for each time interval between pairs of change points. However, this approach relies on a greedy partitioning scheme for determining the change points and the corresponding FC states.

In this paper, we propose a two step approach to track whole brain dynamic connectivity during an event related potential (ERP) recording. As all subjects perform the same task and respond to the same stimulus, we assume that a relatively similar network structure exists for all subjects, and our analysis is tuned to first identify the change points in time where significant changes to the network structure occur across all subjects. We quantify these changes through a low-rank tensor approximation to the FCNs constructed across subjects and subspace distance metrics. Once the change points are detected, the network state for the time intervals between change points is assumed to be stationary and common across subjects, and summarized through tensor-matrix projections across subjects and time. The key contributions of the proposed work are threefold. First, unlike most of the current work, we determine time intervals during which the FCN has a common quasi-stationary pattern across time and subjects. Second, most of the current work reduces the high-dimensionality of the FCNs by vectorizing the connectivity matrices into long vectors before identifying the network

states. This approach may lose some important relationships between the nodes and does not preserve the topological structure of the network. In the proposed work, we address this problem by keeping the network structure of FCNs intact by using tensor representations. Through tensor representation, we can capture the variability common to all subjects across time. Finally, the proposed approach is flexible for extensions to higher order representations such as time- and frequency-dependent network states across time.

II. BACKGROUND

A. Tensor Subspace Analysis

Linear data models such as PCA [48] and Independent Component Analysis [49] are widely used for the decomposition of matrices. Depending on the criteria, different types of basis vectors are extracted and appropriate lower dimensional features are obtained through projection. Multiway data analysis extends these linear methods to capture multilinear structures and underlying correlations in higher order datasets, also known as tensors. Several tensor decomposition methods such as canonical decomposition, a.k.a parallel factor analysis (PARAFAC), and Tucker decomposition have been proposed [50]–[52].

Let $\mathcal{X} \in \mathbb{R}^{m_1 \times m_2 \times m_3}$ be a three-way tensor, where m_i is the dimension along the ith mode, and x_{i_1,i_2,i_3} denotes the (i_1, i_2, i_3)th element of the tensor $\mathcal{X}$. The n-rank of the tensor $\mathcal{X}$ is the column rank of unfolding the tensor along its nth mode [53]. A subtensor is defined by fixing the indices along one or more modes. For instance, when an index of a three-way tensor is fixed in one mode, the resulting matrix is called a slice.

The Tucker decomposition is a higher order generalization of singular value decomposition (SVD) for tensors [50]. The full Tucker decomposition of the three-way tensor $\mathcal{X}$ can be expressed as follows:

$$\mathcal{X} = \mathcal{C} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \times_3 \mathbf{U}^{(3)} + \mathcal{E},$$
$$\mathcal{X} = \sum_{i_1,i_2,i_3} \mathcal{C}_{i_1,i_2,i_3} \left(\mathbf{u}_{i_1}^{(1)} \circ \mathbf{u}_{i_2}^{(2)} \circ \mathbf{u}_{i_3}^{(3)} \right) + \mathcal{E}_{i_1,i_2,i_3} \quad (1)$$

where $\mathcal{C} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is the core tensor, where $n_i \leq m_i$ and $\mathcal{E} \in \mathbb{R}^{m_1 \times m_2 \times m_3}$ is the residual error, which represents the part of data which is not described by the structural part. $\times_k$ is the product of a tensor and a matrix along mode-k. $\mathbf{U}^{(1)} \in \mathbb{R}^{m_1 \times n_1}$, $\mathbf{U}^{(2)} \in \mathbb{R}^{m_2 \times n_2}$, and $\mathbf{U}^{(3)} \in \mathbb{R}^{m_3 \times n_3}$ are the matrix of the left singular vectors of the tensor unfolded along the ith mode, called component matrices. Therefore, the kth singular vector along mode i, $\mathbf{u}_k^{(i)}$ is the kth column of the component matrix $\mathbf{U}^{(i)}$. Various algorithms such as alternating least squares (ALS) [54], Simultaneous matrix diagonalizations [55], Simultaneous generalized Schur decomposition [56], etc., have been proposed to calculate these models. The most popular method for computing Tucker decomposition is the ALS method. The algorithm we use in this paper is called higher-order orthogonal iteration (HOOI), which is the extension of ALS for n-way tensors [52].

While the Tucker decomposition cannot determine the component matrices uniquely like PARAFAC, its component matrices are orthogonal. The orthogonality property of the Tucker

decomposition makes it an appropriate decomposition for lower dimensional subspace projection and the core array represents the interaction between factors from different modes [50]. The tensor $\mathcal{X} \in \mathbb{R}^{m_1 \times m_2 \times m_3}$ can be projected to a lower dimensional tensor $\mathcal{Y} \in \mathbb{R}^{r_1 \times r_2 \times r_3}$ using the projection matrices $\tilde{\mathbf{U}}^{(i)} \in \mathbb{R}^{m_i \times r_i}$ [57], as:

$$\mathcal{Y} = \mathcal{X} \times_1 \tilde{\mathbf{U}}^{(1)\dagger} \times_2 \tilde{\mathbf{U}}^{(2)\dagger} \times_3 \tilde{\mathbf{U}}^{(3)\dagger} \quad (2)$$

where $\dagger$ is the matrix transpose operator, and the projection matrix $\tilde{\mathbf{U}}^{(i)}$ is formed by taking the first r_i columns of $\mathbf{U}^{(i)}$, where $r_i \leq n_i$.

B. Time-Varying Network Construction

After recording the EEG data, they are preprocessed by the spherical spline current source density (CSD) waveforms to sharpen ERP scalp topographies and reduce volume conduction [58]. CSD performs spatial filtering of each channel taking into account its immediate neighbors. CSD reflects the local regional activity around that channel rather than activity in anatomically specified (parcellated) brain regions. The CSD has fewer assumptions than many inverse transforms, attenuates volume conduction, and represents independent sources near the cortical surface [59]. Then, the functional connectivity brain networks are constructed from the filtered multichannel EEG data where the nodes correspond to the different channels, reflecting the activity around that channel, and the edges to the connectivity between these regions.

In this paper, we quantify the connectivity using a recently introduced phase synchrony measure based on reduced interference distribution (RID)-Rihaczek distribution [60]. The first step in quantifying phase synchrony is to estimate the time and frequency-dependent phase $\Phi_i(t, \omega)$ of a signal s_i, $\arg\left[\frac{C_i(t,\omega)}{|C_i(t,\omega)|}\right]$, where $C_i(t, \omega)$ is the complex RID-Rihaczek distribution[1]:

$$C_i(t, \omega)$$
$$= \iint \underbrace{\exp\left(-\frac{(\theta\tau)^2}{\sigma}\right)}_{\text{Choi-Williams kernel}} \underbrace{\exp\left(j\frac{\theta\tau}{2}\right)}_{\text{Rihaczek kernel}} A_i(\theta, \tau) e^{-j(\theta t + \tau\omega)} d\tau d\theta$$

$$(3)$$

and $A_i(\theta, \tau) = \int s_i(u + \frac{\tau}{2}) s_i^*(u - \frac{\tau}{2}) e^{j\theta u} du$ is the ambiguity function of the signal s_i. The phase synchrony between channels i and j at time t and frequency ω is computed using phase locking value (PLV):

$$PLV_{(i,j)}(t, \omega) = \frac{1}{L} \left| \sum_{k=1}^{L} \exp\left(j\Phi_{(i,j)}^k(t, \omega)\right) \right| \quad (4)$$

where L is the number of trials in the ERP measurements and $\Phi_{(i,j)}^k(t, \omega) = |\Phi_i^k(t, \omega) - \Phi_j^k(t, \omega)|$ is the phase difference between the two channels for the kth trial. The single-trial phase differences, represented by complex numbers, are averaged across trials. The absolute value of this complex average

---

[1]Details of the RID-Rihaczek distribution and the corresponding synchrony measure are given in [60].

yields the magnitude value that defines the PLV. A PLV of 0 indicates completely uniform random phase angle differences across trials, whereas PLV of 1 indicates completely consistent phase angle differences across trials. This measure quantifies the extent to which oscillation phases are similar across different electrodes at each time-frequency point.

Once the pairwise synchrony values are computed at each time and frequency point, we can construct a <mark>time-varying connectivity graph</mark> $\{\mathbf{G}(t); \forall t \in \{t = 1, 2, \ldots, T\}\}$ with

$$G_{(i,j)}(t) = \frac{1}{\Omega} \sum_{\omega=\omega_a}^{\omega_b} PLV_{(i,j)}(t, \omega) \quad (5)$$

where $G_{(i,j)}(t) \in [0, 1]$ represents the connectivity strength between the nodes i and j within the frequency band of interest, $[\omega_a, \omega_b]$, and Ω is the number of frequency bins in that band.

$\mathbf{G}_s(t)$ is the adjacency matrix at time step t for the sth subject, where $G_{s,(i,j)}(t) \forall i, j \in \{1, 2, \ldots, N\}$ indicates the strength of the connection between nodes i and j. It is assumed that <mark>all subjects perform the same task at the same time</mark>, so the collection of S subjects at time t forms a three-way tensor $\mathcal{G}(t) \in \mathbb{R}^{N \times N \times S}$, where $\mathbf{G}_s(t) = \mathcal{G}_{i_3=s}(t)$.

III. CHANGE-POINT DETECTION

Assume that the sequence $\mathcal{G}(t)$ contains $\tau + 1$ functional connectivity states [32], [41], [61], with change points located at $\{t_1, t_2, \ldots, t_\tau\}$, where the location and the number of change points are not known *a priori*. In order to detect the boundaries of functional connectivity states t_k, where $k \in \{1, 2, \ldots, \tau\}$ stands for the kth change point, we calculate the distance between the lower dimensional subspaces that best approximate the connectivity graphs along the three modes of $\mathcal{G}(t)$ for $t = 1, 2, \ldots, T$. This approach is motivated by the fact that the most real data lie in a low-dimensional subpace or a union of subspaces [62].

In order to find the change points which are consistent among all subjects, we use a bootstrapping approach as follows [63]. $\bar{S}$ subjects are chosen randomly without replacement [see Fig. 1(b)]. In order to find the optimal low-rank of the the tensor $\mathcal{G}(t) \in \mathbb{R}^{N \times N \times \bar{S}}$, $(\hat{n}(t), \hat{n}(t), \hat{s}(t))$, the tensor is fully decomposed for the selected subjects [see Fig. 1(c)]. Using the convex hill algorithm, the optimal low-rank is calculated from the fully decomposed tensor [see Fig. 1(d)]. The convex hull algorithm evaluates the goodness of fit between the original tensor and its low-rank approximation for different combinations of ranks and selects the optimal value numerically [64]. Since the adjacency matrix $\mathbf{G}_s(t)$ is symmetric, the first and second modes are forced to have the same rank $\hat{n}(t)$ [65], [66].

Given the ranks of the tensor along each mode $(\hat{n}(t), \hat{n}(t), \hat{s}(t))$, we used HOOI to compute a low-rank approximation of $\mathcal{G}(t)$, denoted by $\hat{\mathcal{G}}(t)$, with rank $(\hat{n}(t), \hat{n}(t), \hat{s}(t))$ [52]. The three-way tensor $\hat{\mathcal{G}}(t)$ at time step t is decomposed using Tucker decomposition as follows:

$$\hat{\mathcal{G}}(t) = \mathcal{C}(t) \times_1 \mathbf{U}^{(1)}(t) \times_2 \mathbf{U}^{(2)}(t) \times_3 \mathbf{U}^{(3)}(t) \quad (6)$$

where $\mathcal{C}(t) \in \mathbb{R}^{\hat{n}(t) \times \hat{n}(t) \times \hat{s}(t)}$ is the core tensor, and $\mathbf{U}^{(1)} \in \mathbb{R}^{N \times \hat{n}(t)}$, $\mathbf{U}^{(2)} \in \mathbb{R}^{N \times \hat{n}(t)}$, and $\mathbf{U}^{(3)} \in \mathbb{R}^{\bar{S} \times \hat{s}(t)}$ are the matrix of the left singular vectors. The bases

$\mathbf{u}_i^{(1)}(t)$, $i \in \{1, 2, \ldots, \hat{n}(t)\}$, $\mathbf{u}_j^{(2)}(t)$; $j \in \{1, 2, \ldots, \hat{n}(t)\}$, and $\mathbf{u}_k^{(3)}(t)$, $k \in \{1, 2, \ldots, \hat{s}(t)\}$ span the lower dimensional subspaces. Note that $\mathbf{U}^{(1)}(t) = \mathbf{U}^{(2)}(t)$ in the case of symmetric connectivity and we will use this equality in the equations that follow to simplify the representation.

Once $\mathbf{U}^{(1)}(t) \forall t \in \{1, 2, \ldots, T\}$ are determined, the distance between subspaces of the first mode of $\mathcal{G}(t)$, $\forall t \in \{1, 2, \ldots, T\}$ need to be calculated as we are interested in detecting the changes along the connectivity mode. The subspace distance is calculated between consecutive time points using $\mathbf{U}^{(1)}(t)$. Since subspaces of the first mode of $\mathcal{G}(t)$ and $\mathcal{G}(t + 1)$ may be of different dimensions $\hat{n}(t)$ and $\hat{n}(t + 1)$, ordinary subspace distance measures such as Euclidean distance cannot be used. Thus, we quantify the subspace distance using the <mark>Grassmann distance</mark> [67] between $(\mathcal{G}(t), \mathcal{G}(t + 1))$ $\forall t \in \{1, 2, \ldots, T - 1\}$ along the first mode, as follows:

$$d(t) = \frac{1}{\sqrt{r}} \sqrt{\sum_{k=1}^{r} \theta_k^2 \left(\mathbf{U}^{(1)}(t), \mathbf{U}^{(1)}(t + 1)\right)} \quad (7)$$

where $r = \min(\hat{n}(t), \hat{n}(t + 1))$ and $\theta_k \left(\mathbf{U}^{(1)}(t), \mathbf{U}^{(1)}(t + 1)\right)$ is the kth principal angle between subspaces of the first mode of $\mathcal{G}(t)$ and $\mathcal{G}(t + 1)$ (see Appendix A for more details) [67].

In order to decide whether $d(t)$ is a change point, the subspace distance $d(t)$ needs to be converted into a binary value $v(t)$ through thresholding, as follows:

$$v(t) = \begin{cases} 1 & \text{if} \quad d(t) \geq \mu_d + 2\sigma_d \\ 0 & \text{if} \quad d(t) < \mu_d + 2\sigma_d \end{cases}$$

$$\mu_d = \frac{1}{T - 1} \sum_{t=1}^{T} d(t)$$

$$\sigma_d = \sqrt{\frac{1}{T - 2} \sum_{t=1}^{T} (d(t) - \mu_d)^2} \quad (8)$$

where the threshold is selected as $\mu + 2\sigma$ since that corresponds to the 95% significance level based on the observation that the distances are distributed as Gaussian. <mark>This process is repeated B times for different random samplings across the subject mode in order to find change points which are consistent among all different random samplings.</mark> Thus, the binary time series $v(t)$ across B samplings are summed up. The final set of change points are determined as the time points when this sum is at a local maxima. Fig. 1 illustrates the steps of change-point detection algorithm.

IV. FUNCTIONAL CONNECTIVITY STATE SUMMARIZATION

After identifying the time boundaries of the functional connectivity states, the goal is to summarize the functional connectivity patterns across time and subjects within the time interval to derive the state connectome. Previously, this was commonly addressed by averaging the edges over the time interval [41], [68]. This method has the drawback of emphasizing all of the edges equally and resulting in very dense network representations.
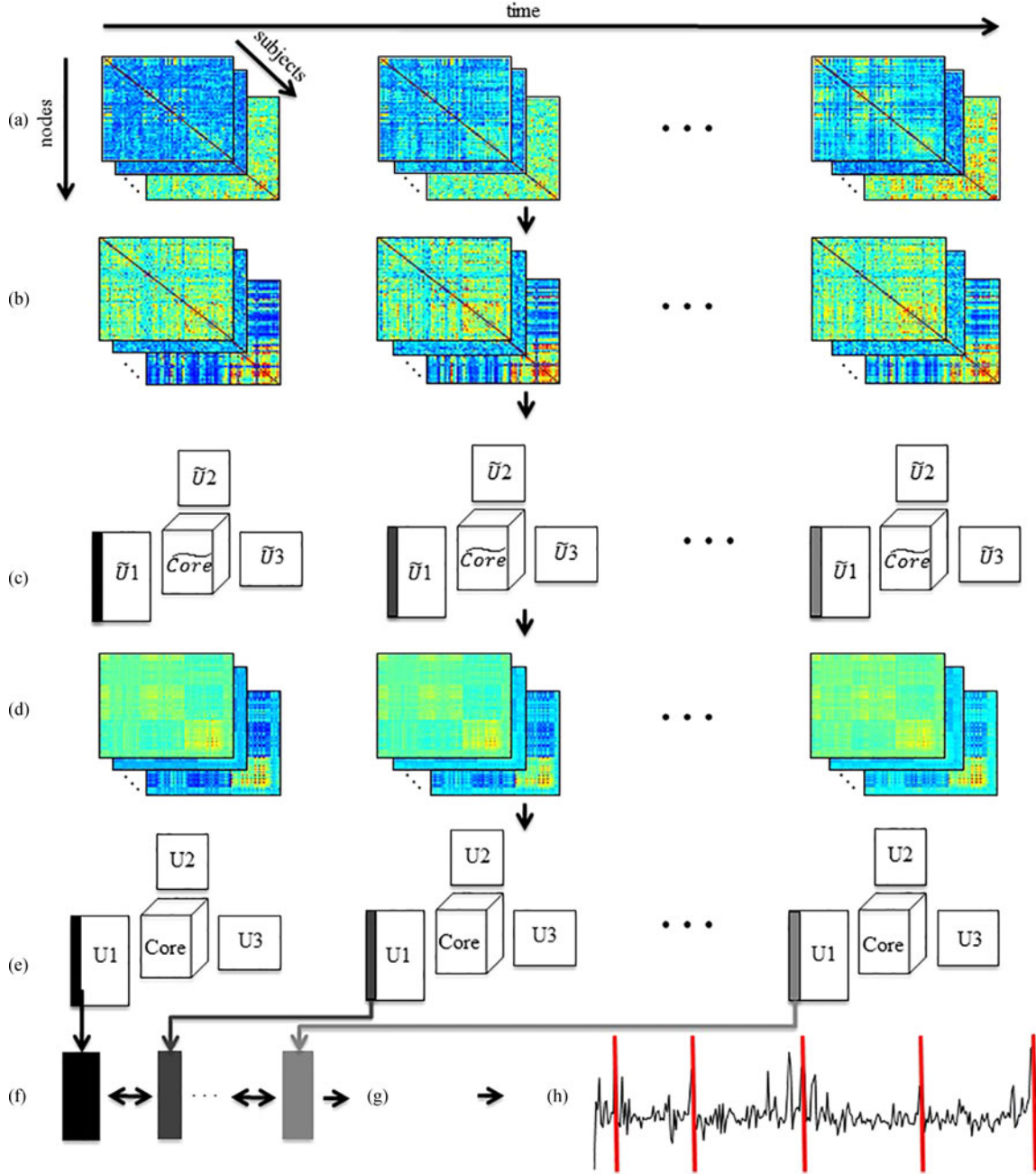
Fig. 1. Flowchart of the change-point detection algorithm described in Section III, (a) original FC tensors formed across all subjects at each time point, (b) random subset of subjects is selected for bootstrapping, (c) full Tucker decomposition of FC tensors corresponding to the selected subjects is performed to determine the best low-rank approximation along each mode, (d) low-rank approximation of FC tensors using convex hull approximation, (e) Tucker decomposition of the low-rank approximation obtained from (d), (f) subspace distance calculation between consecutive time steps using the basis elements (columns of) U_1, (g) subspace distances thresholding, and (h) change point detection.

The functional connectivity states are assumed to be quasi-stationary, which means that the corresponding connectivity patterns do not change significantly within the network states. Therefore, the entire network state can be represented by a single adjacency matrix demonstrating the most significant interactions among various brain regions.

A. Functional Connectivity Summarization

1) Subject Summarization: Since all subjects perform the same task at the same time, we assume that the connectivity networks across subjects are consistent for a given time interval.

Given the time interval (t_k, t_{k+1}), all three-way tensors $\mathcal{G}(t)$ in this time interval are individually fully decomposed using Tucker decomposition as follows:

$$\mathcal{G}(t) = \mathcal{C}(t) \times_1 \mathbf{U}^{(1)}(t) \times_2 \mathbf{U}^{(1)}(t) \times_3 \mathbf{U}^{(3)}(t) \qquad (9)$$

where $\mathcal{C}(t)$ is the core tensor, and $\|\mathcal{C}_{i_l=1}(t)\|^2 \geq \|\mathcal{C}_{i_l=2}(t)\|^2 \geq \cdots \geq \|\mathcal{C}_{i_l=N}(t)\|^2$ for $l = 1$ and 2, and $\|\mathcal{C}_{i_3=1}(t)\|^2 \geq \|\mathcal{C}_{i_3=2}(t)\|^2 \geq \cdots \geq \|\mathcal{C}_{i_3=S}(t)\|^2$ are the singular values along the first, second, and third mode, respectively.

We define $\zeta(t) = \mathcal{C}(t) \times_1 \mathbf{U}^{(1)}(t) \times_2 \mathbf{U}^{(1)}(t)$, which leads to an alternative expression of Tucker decomposition

as $\mathcal{G}(t) = \zeta(t) \times_3 \mathbf{U}^{(3)}(t)$. Due to the orthogonality of $\mathbf{U}^{(1)}(t)$, we can conclude from $\|\mathcal{C}_{i_3=1}(t)\|^2 \geq \|\mathcal{C}_{i_3=2}(t)\|^2 \geq \cdots \geq \|\mathcal{C}_{i_3=S}(t)\|^2$ that $\|\zeta_{i_3=1}(t)\|^2 \geq \|\zeta_{i_3=2}(t)\|^2 \geq \cdots \geq \|\zeta_{i_3=S}(t)\|^2$ [69]. Thus, the subtensor

$$\gamma(t) = \zeta_{i_3=1}(t) = \sum_{s=1}^{S} U_{s,1}^{(3)}(t)\mathbf{G}_s(t)$$

where $\gamma(t) \in \mathbb{R}^{N \times N}$ is the factor that captures most of the energy of the activation patterns across subjects at time t. Similarly, $\zeta_{i_3=2}(t)$ captures the second largest amount of energy across all subjects at time t [69], [70]. The subtensor $\gamma(t)$ is the common functional connectivity pattern across all subjects at the time step t.

2) _Time Summarization:_ After the connectivity information across all subjects is summarized and represented by a single adjacency matrix, the network state $\gamma(t) \, \forall t \in \{t_k, t_k + 1, \ldots, t_{k+1}\}$ needs to be summarized across time mode within the detected functional connectivity state boundaries in order to derive the state connectome.

First, we construct a three-way tensor $\mathbf{\Gamma} \in \mathbb{R}^{N \times N \times (t_{k+1}-t_k+1)}$ from the set of matrices $\gamma(t) \in \mathbb{R}^{N \times N}$, where the third mode is time. The tensor $\mathbf{\Gamma}$ is fully decomposed using Tucker decomposition as follows:

$$\mathbf{\Gamma} = \vartheta \times_1 \bar{\mathbf{U}}^{(1)} \times_2 \bar{\mathbf{U}}^{(1)} \times_3 \bar{\mathbf{U}}^{(3)} = \bar{\zeta} \times_3 \bar{\mathbf{U}}^{(3)}. \qquad (10)$$

The subtensor $\eta = \bar{\zeta}_{i_3=1} = \sum_{t=t_k}^{t_{k+1}} \bar{U}_{t,1}^{(3)} \mathbf{\Gamma}_{i_3=t}$, where $\eta \in \mathbb{R}^{N \times N}$, captures the largest amount of energy across all time steps. Thus, it can be used as the common functional connectivity state across all subjects and time steps within the time interval (t_k, t_{k+1}).

B. Significance Testing

The subtensor η represents the summarization of functional connectivity across subjects and time within the time interval (t_k, t_{k+1}). In order to characterize the FC within a state more concisely, we need to determine the significant edges of η. Thus, a hypothesis testing on η obtained for each time interval is required in order to identify the significant edges. Since the histogram of the edge values in η appear to be bell shape and symmetric, we assume a Gaussian distribution for the edge values in η (see Fig. 7). This assumption can be validated using Kolmogorov–Smirnov test [71], [72], where the null hypothesis is that the edge values in η are drawn from a Gaussian distribution family. A z-test can then be used on the edges of η to determine the most significant edges. Fig. 2 shows an overview of the proposed summarization algorithm.

C. Computational Complexity

The computational complexity of the change point detection algorithm depends mostly on the complexity of HOOI since HOOI requires the decomposition of the tensor with various ranks in order to find the optimal low-rank. Let us assume that HOOI decomposes the tensor with I different sets of ranks. The complexity of each decomposition depends on the selected

rank. The full-rank decomposition has the highest computational complexity, which is equal to $\mathcal{O}((\frac{N+N+S}{3})^4)$ [73]. Assuming full-rank decomposition for all HOOI iterations (the worst case), the computational complexity of HOOI is at most equal to $\mathcal{O}(I(\frac{N+N+S}{3})^4)$. Even though HOOI procedure is performed for every time step, HOOI procedure for all time steps can be computed in parallel since HOOI for each time step is independent of another. Similarly, the computational complexity of the full Tucker decomposition and the $(\hat{n}, \hat{n}, \hat{s})$-rank decompositions are $\mathcal{O}((\frac{N+N+S}{3})^4)$ and $\mathcal{O}((\frac{\hat{n}+\hat{n}+\hat{s}}{3})^4)$, respectively. Thus, the total computational complexity of the change point detection algorithm step is $\mathcal{O}((I + 1)(\frac{N+N+S}{3})^4 + (\frac{\hat{n}+\hat{n}+\hat{s}}{3})^4)$.

The computational complexity of FC state representation consists of two parts, computational complexity of summarization across subjects and time. Each summarization step includes the computational complexity of the Tucker decomposition and that of the projection step. The computational complexity of the multiplication of a matrix with size $d_1 \times d_2$ by a vector with size $d_2 \times 1$ is $\mathcal{O}(d_1 d_2)$ [74]. The computational complexity of Tucker decomposition across subjects at each time step is $\mathcal{O}(S(\frac{N+N+S}{3})^4)$, and the computational complexity of the projection of subjects onto a single matrix at each time step is $\mathcal{O}(S^2 N^2)$. The total computational complexity for the subject summarization step is $\mathcal{O}((\frac{N+N+S}{3})^4 + S^2 N^2)$ since the summarization of all subjects are independent and can be done in parallel. Similarly, the total computational complexity of summarization across time is $\mathcal{O}((\frac{N+N+\tau}{3})^4 + \tau^2 N^2)$, where $\tau = t_{k+1} - t_k + 1$. Therefore, the total computational complexity of the summarization step is $\mathcal{O}((\frac{N+N+S}{3})^4 + (\frac{N+N+\tau}{3})^4 + (\tau^2 + S^2)N^2)$.

V. Experimental Results

In this section, the results of the proposed method both on simulated networks and EEG FCNs are provided to evaluate the performance of the proposed method.

A. Simulated Network

In this section, we generate a simulated dynamic network that resembles the data structure of our EEG FCNs to evaluate the effectiveness of the proposed approach. A tensor corresponding to a dynamic network with $N = 64$ nodes, $S = 32$ subjects, and $T = 128$ time steps is generated. The change points are fixed at $t_1 = 41$ and $t_2 = 81$.

Within the time interval $[1, 40]$, all graphs at each time point and across all subjects are generated with a two cluster structure. The first cluster consists of nodes $(1 - 48)$, while nodes $(49 - 64)$ belong to the second cluster. Intracluster edges have higher mean edge value and lower variance compared to intercluster edges. The weights of the intracluster edges are selected randomly from $\mathcal{N}(0.2, 0.01)$ while the weights of the intercluster edges are chosen randomly from $\mathcal{N}(0.1, 0.02)$. Out of 1128 edges within the first cluster, 500 significant edges are selected randomly using R-MAT algorithm [75]. In order for the selected edges to be significantly important, the weights are selected from $\mathcal{N}(\mu, 0.01)$, where $\mu \in (0.2, 0.6)$ and added
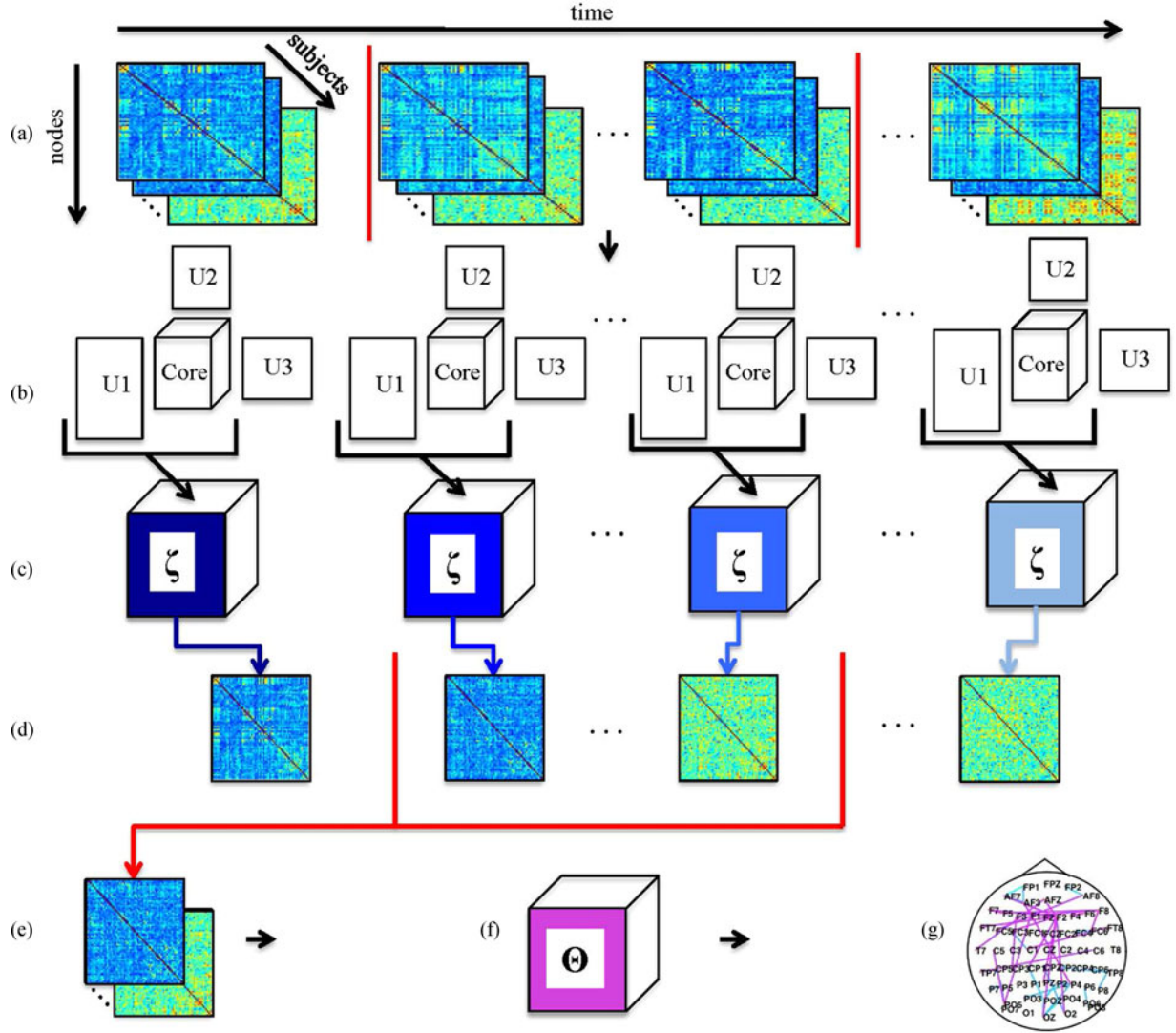
Fig. 2. Functional connectivity state summarization algorithm flowchart described in Section IV-A. (a) Original connections and event intervals, (b) Tucker decomposition, (c) Consolidation: define η, (d) Take first slice of η, (e) Collect slices in an event interval, (f) Repeat Tucker decomposition and consolidation, (g) Summarize network usingh first slice of Θ and significant testing.

to the intracluster edge values. Similarly, 50 out of 120 edges from the second cluster are chosen randomly using R-MAT algorithm, and their weights are determined by adding random values from $\mathcal{N}(\mu, 0.01)$ with the same μ value to the intracluster edges.

For the time interval $[41, 80]$, nodes $(1 - 16)$ at each time step and for all the subjects belong to the first cluster while nodes $(17 - 64)$ belong to the second cluster. The weights of all the intracluster edges are selected randomly from $\mathcal{N}(0.2, 0.01)$ while the weights of the intercluster edges are generated from $\mathcal{N}(0.1, 0.02)$. A total of 50 significant edges within the first cluster and 500 significant edges within the second cluster are selected randomly using R-MAT algorithm. The weights of the significant edges for all subjects at each time step are formed by selecting random values from the distribution $\mathcal{N}(\mu, 0.01)$, where $\mu \in (0.2, 0.6)$, and adding them to the intracluster edges. Graphs for all subjects at each time step within the time interval

$[81, 128]$ are generated similar to the graphs within the time interval $[1, 40]$.

To evaluate the performance of the proposed algorithm under different conditions, five experiments with μ values in the range $(0.2, 0.6)$ are run. Each experiment is repeated 400 times. The signal to noise ratio (SNR) is calculated for different values of μ, where signal power is defined as the total power of the significant edges ($\mathcal{N}(\mu, 0.01)$), and the noise power is defined as the total power of inter and intracluster edges. Let $n_{s,(i,j)}(t)$ be the background connectivity strength of subject s between nodes i and j at time t, and $n_{s,(i,j)}(t) + f_{s,(i,j)}(t)$ be the connectivity of the subject s between nodes i and j at time t, where $f_{s,(i,j)}(t) \sim \mathcal{N}(\mu, 0.01)$. Then, SNR is defined as follows:

$$SNR = \frac{\frac{1}{STN^2} \sum_{t=1}^{T} \sum_{s=1}^{S} \sum_{i,j=1}^{N} f_{s,(i,j)}^2(t)}{\frac{1}{STN^2} \sum_{t=1}^{T} \sum_{s=1}^{S} \sum_{i,j=1}^{N} n_{s,(i,j)}^2(t)}$$
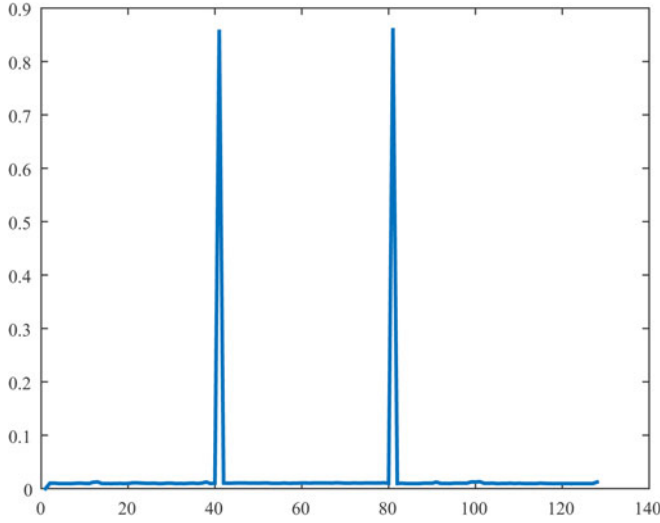
Fig. 3.   Average subspace distance of the simulated network for $\mu = 0.90$ with two peaks at $t = 41$ and $t = 81$.

TABLE I
ERROR PROBABILITY FOR SIMULATED DYNAMIC NETWORKS DESCRIBED IN SECTION V-A

SNR	−10.63 dB (%)	−7.13 dB (%)	−4.66 dB (%)	−2.71 dB (%)	−1.11 dB (%)
Type *I* (miss)	32.8	17	10.8	10.2	9.9
Type *II* (false alarm)	3.53	2.41	1.64	1.46	1.37

where S is the total number of subjects, T is the total number of time steps, and N is the number nodes.

1) Functional Connectivity State Detection: The change point detection algorithm is run on all 400 simulations of this dynamic network model. The binary vectors $v(t)$ across 400 simulations are averaged for each μ. The peaks of the average subspace distance for each μ indicate the change points as seen in Fig. 3.

2) Functional Connectivity State Representation: Each of the three time intervals that was detected by the subspace distance measure were summarized individually across 400 simulations. z-test with a p-value of 0.95 is performed on the summarized graphs. Since the weights of the edges are not recovered during summarization, there is no ground-truth to compare η to. Therefore, we compare only the location of the detected significant edges within each time interval, and calculate the probability of type *I* (false positive) and *II* (false negative) errors averaged across each experimental condition, (μ, σ), i.e., different SNRs. The probability of type *I* refers to the number of those significant edges which are not detected to the total number of significant edges. The probability of type *II* is the number of detected edges which belong to the background to the total number of normal edges. Table I summarizes the results for each SNR value. It can be seen that as SNR increases, i.e., the strength of the significant edges, the probability of both

type *I* and *II* errors decreases as expected since the significant edges are easier to detect with large values of SNR. In practice, most FCNs are expected to have a SNR in the range of 0 dB and as such will result in Type I and II errors below 10%.

B. EEG Data

The proposed framework is applied to a set of EEG data containing the ERN. The ERN is a brain potential response that occurs following performance errors in a speeded reaction time task usually 25–75 ms after the response [76].

EEG data from 62-channels was collected in accordance with the 10/20 system on a Neuroscan Synamps2 system (Neuroscan, Inc.) at a sampling rate of 128 Hz. This study included 91 undergraduate students (34 male) from the University of Minnesota. Full methodological details of the recording are available in a previous paper [76]. The task was a common speeded-response letter (H/S) flanker, where error response-locked trials from each subject were utilized.

Previous works [77], [78] indicate that there is an increased phase synchrony associated with ERN for the theta frequency band ((4, 7) Hz) and ERN time window ((25, 75) ms) for Error responses compared to Correct responses [78]. Moreover, this increased synchrony is observed between the electrodes corresponding to the LPFC and medial prefrontal cortex.

The time-varying connectivity matrices in this study are constructed by computing the pairwise PLV over an average of 24 trials between 62 channels in the theta frequency band for all time (2 s corresponding to 256 time samples) and all subjects using (4), resulting in $\mathcal{G}(t) \in \mathbb{R}^{62 \times 62 \times 91}$ for the error response.

1) Functional Connectivity Change Point Detection: We selected $\bar{S} = 40$ subjects out of 91 subjects randomly without replacement to calculate the subspace distance between consecutive time points and detect change points. This procedure is repeated 106 times. This process takes about 38 s to be completed. Fig. 4(b) shows the number of change points at each time point across 106 random samplings. By choosing the peaks of this time series, change points are detected to be at $\{-859, -656, -180, -39, 55, 281, 469, 664\}$ ms. Right before ERN time interval, $(-39, 55)$ ms, the functional connectivity network is not stable due to the differences in response time across subjects. This leads to the observation of multiple peaks right before the 0 time point which we do not consider as valid change points.

In order to show the effectiveness of tensor-based methods compared to the traditional matrix-based methods, the upper triangle part of $\mathbf{G}_s(t)$, $s \in \{1, 2, \ldots, S\}$ is vectorized and concatenated across subjects to form a matrix. This process is repeated for all time steps so that T subject-connectivity matrices of size 1953×91 are obtained similar to [43]. SVD is performed on the matrices at each time step similar to the tensor approach without subject sampling. An optimal rank at each time point is determined using the convex hull algorithm. Given the rank at each time, the Grassmann distance between consecutive time steps is computed. Fig. 5 shows the Grassman distance between consecutive time steps based on SVD decomposition of the vectorized adjacency matrices, and based on the tensor projection
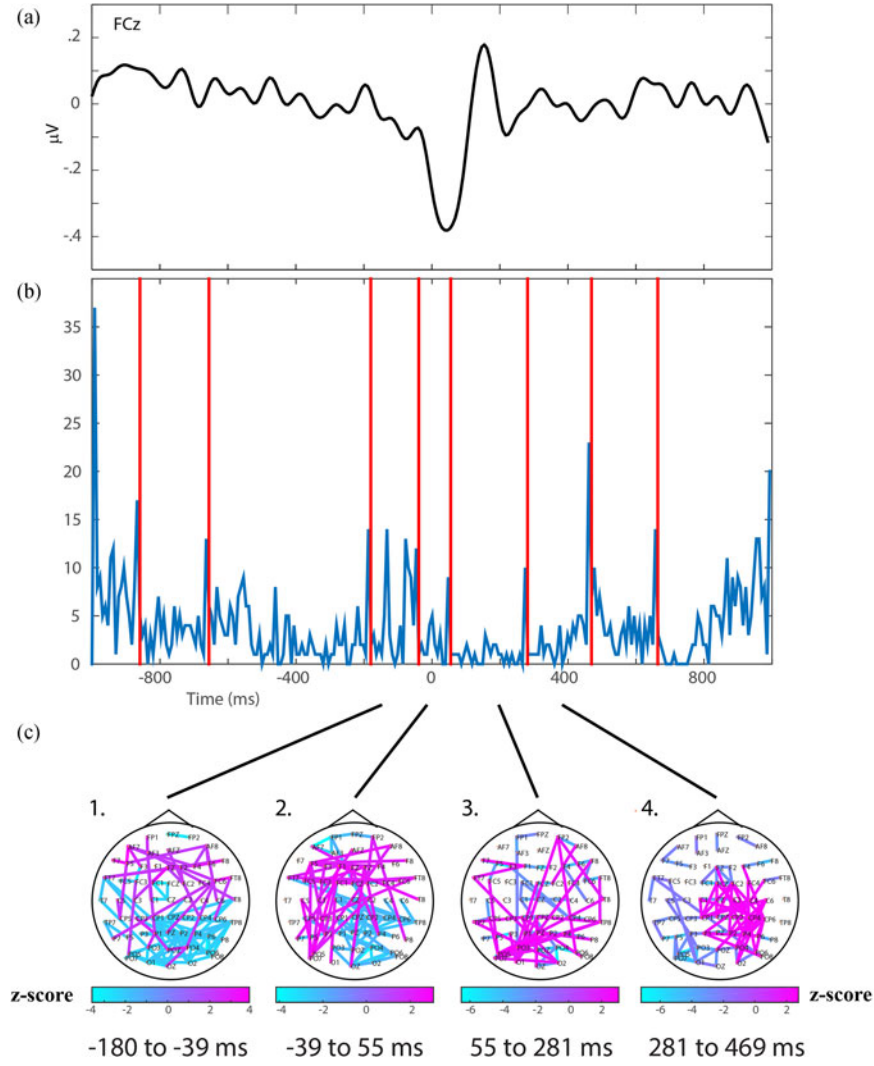
Fig. 4. (a) Waveform for error trials across all subjects, 0 ms marks the time of the button-press incorrect response. (b) The binary vector resulted from the subspace distance between the projection operators at consecutive time points, calculated using the Grassman distance, where red lines mark the identified change points. (c) The most significant edges of η with $p = 0.95$ for time intervals 1. $(-180, -39)$ ms, 2. $(-39, 55)$ ms, 3. $(55, 281)$ ms, 4. $(281, 469)$ ms.

method.[2] In the graph corresponding to SVD method, a few change points are obvious, among them, the one close to zero corresponds to the beginning of ERN. However, other change points such as the end of ERN are not detectable. Although the scales of SVD and tensor-based methods are different, it can be seen that the tensor decomposition is more sensitive to the changes in the network structure with more apparent change points before and after the response. Both methods yield similar peaks around ERN which verifies the accuracy of the proposed tensor-based method.

2) Functional Connectivity State Representation [see Fig. 4(c)]: Once change points are determined, tensor projection along subject and time modes for all nine time intervals are performed to obtain the FC-states $(\eta_1, \eta_2, \ldots, \eta_9)$. In

order to detect the significant edges within each interval, the upper triangular part of η_i is converted to a vector with $\binom{63}{2} = 1953$ entries, and Gaussian distributions with different mean and variances are fitted to the vectors. $\mathcal{N}_1(0.22, 0.01)$, $\mathcal{N}_2(0.32, 0.01)$, $\mathcal{N}_3(0.07, 0.01)$, $\mathcal{N}_4(-0.44, 0.01)$, $\mathcal{N}_5(-0.08, 0.01)$, $\mathcal{N}_6(0.24, 0.01)$, $\mathcal{N}_7(-0.40, 0.02)$, $\mathcal{N}_8(-0.40, 0.01)$, and $\mathcal{N}_9(-0.25, 0.02)$ are the Gaussian distributions fitted to $(\eta_1, \eta_2, \ldots, \eta_9)$, respectively. The implementation of the summarization algorithm is fast, for example, it takes approximately 6 s to find the summarization of ERN time interval. Fig. 7 shows the histogram of η_5 corresponding to ERN time interval and the fitted Gaussian distribution. Kolmogorov–Smirnov test is used to measure the goodness-of-fit for the assumed Gaussian distributions. Kolmogorov–Smirnov test with $\alpha = 0.01$ does not reject the null hypothesis that the edges follow a Gaussian distribution.

The summarized functional connectivities $(\eta_1, \eta_2, \ldots, \eta_9)$ are normalized with respect to the estimated means and variances. z-test with $\alpha = 0.05$ is performed to determine the signifi-

[2]Note that in this experiment random sampling of subjects is not performed, and these are distances not the binary vectors. That is why the result given here for the tensor projection is different from Fig. 4.
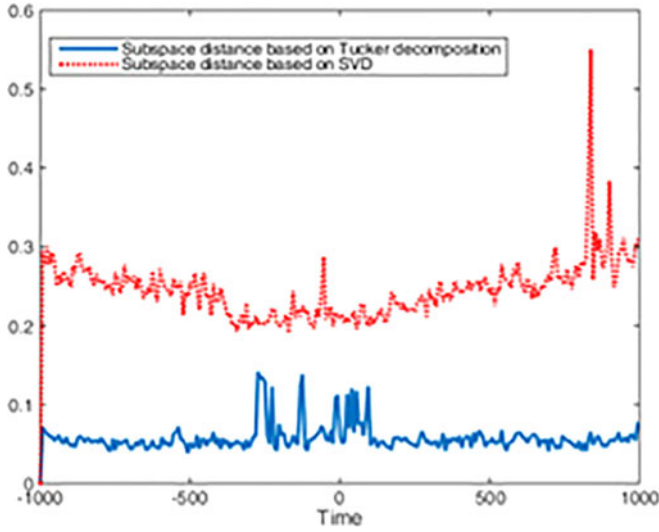
Fig. 5. The Grassmann distance between consecutive time steps of the functional connectivity network calculated based on SVD and tensor projection.
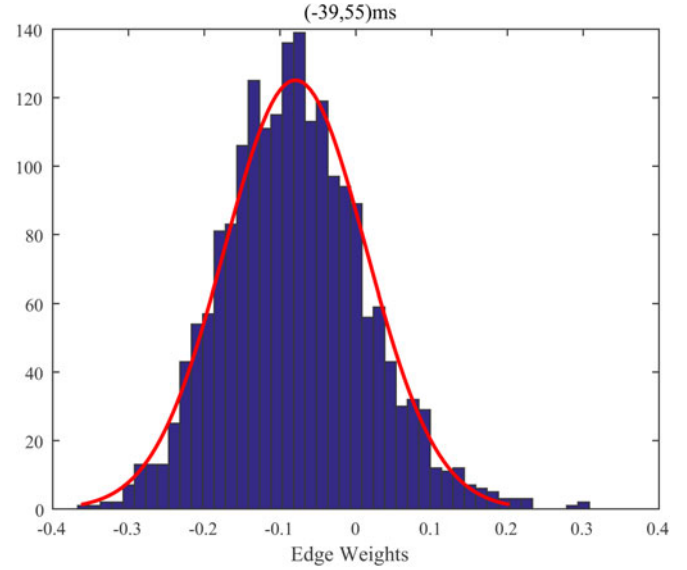

Fig. 7. Histogram of the edges in network state η_5 corresponding to ERN time interval and the fitted Gaussian distribution $\mathcal{N}_2(-0.08, 0.01)$.
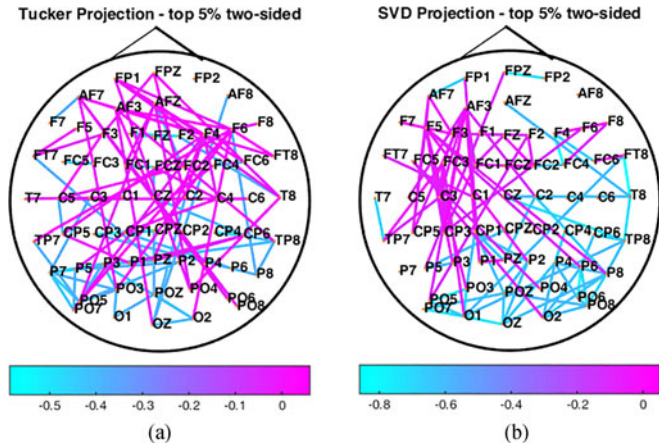

Fig. 6. (a) Top 5% of the projection result using SVD method. (b) Top 5% of the projection result using Tucker decomposition.

cant edges within the nine time intervals. Fig. 4(c) shows the significant edges for four time intervals $(-180, -39)$, $(-39, 55)$, $(55, 281)$, and $(281, 469)$ ms.

Similar to the previous section, the result of the tensor projection method is compared with the result of SVD-based method. However, since the change points obtained from the SVD approach are not exactly the same as the change points from the tensor approach, the summarized networks would correspond to different time intervals and the comparison would not be fair. Therefore, the ERN time interval obtained from the tensor-based method is used for comparison purposes. For the SVD approach, the matrices constructed by vectorizing and stacking the FC matrices of subjects at each time step are projected onto their first singular vectors to summarize the subject information. The resulting vectors at each time step are concatenated across time to form a new matrix. The new matrix is projected onto its first singular vector to summarize the whole information across

time and subject. The two-sided top 5% of edges are shown in Fig. 6(a) and (b) for SVD and tensor-based methods, respectively. Substantial previous research, including several reports from our group, supports connections between medial and LPFC area (especially right-frontal). This pattern of connections is apparent for the tensor-based approach Fig. 6(b), with significant connections centered on electrodes F6, F4. For the SVD-based approach presented in Fig. 6(a), on the other hand, the pattern is inconsistent with previous reporting about activity during the ERN. Most obviously, connections for the SVD are dominated by a substantially left-lateralized pattern of fronto-central connections.

3) Discussion: As apparent in Fig. 4(b), the $(-39, 55)$ ms time window corresponds most directly to the ERN, while the two periods after are associated with post-ERN activity, including the Pe, which is a positive deflection of the waveform following the ERN also known as the error-related positivity. Pe is basically the same as the P300 wave associated with conscious sensations. The corresponding connectivity state tensor topographic maps presented in Fig. 4(c) contain the most significant edges of η with $p = 0.95$. Importantly, because polarity is arbitrary for the connectivity state obtained through tensor projection, the significant tails of the distributions indicate inversely related network subparts. That is, for a given observed network state, the two tails index regions of relative increased versus decreased or inhibited connectivity during that state. As will be described next, two of the states appear to have meaningful activity only on one side. The observed pattern of significant functional connectivities are consistent with observed and hypothesized activity associated with the ERN.

First, in Fig. 4(c-2) $(-39, 55)$ ms, increased connectivity is observed in medial-frontal regions, consistent with widely observed ERN amplitude measures. The engagement of lateral–frontal regions as hypothesized and consistent with previous

works [45], [77], [79], is also apparent. Connections are largest among medial and lateral frontal regions in the $(-39, 55)$ ms network state relative to the other network states, consistent with engagement of these regions during the ERN. Interestingly, while some left-lateralized parietal regions are functionally connected to the frontal regions during the ERN window, a larger central and right lateralized parietal cluster is deactivated during increase in frontal activity. This is consistent with a functional state dominated by frontal and lateral–frontal regions. Next, the pattern for the network state prior to the ERN window [see Fig. 4(c.1)] contains a similar structure, with a difference being that, relative to the ERN window, the density of connections among frontal regions is decreased while connections among parietal regions are increased. This is consistent with the idea that the frontal networks are in the process of becoming functionally integrated leading up to the ERN, where this type of functional integration is maximal. Finally, for the remaining two windows [see Fig. 4(c-3) and (c-4)] only one tail represents a densely connected state of functional connectivity (where the edges from the other tail are more distributed and disconnected, inconsistent with an integrated functional network and potentially spurious). Relative to Fig. 4(c-4.) $(55, 281)$ ms, the main connectivity state is parietally centered, with more sparse connections to frontal regions. This is consistent with a period of dissipation of frontal functional networks that were maximal during the ERN. In Fig. 4(c-4) $(281, 463)$ ms the primary network state continues this shift toward parietal regions. Broadly, the network states identified before, during, and after the ERN are consistent with the hypothesized ERN-related network dynamics, focused on the formation of transient frontal networks during error processing.

VI. Conclusion

In this paper, we introduced a two-step approach to dFC network identification and characterization. The proposed method is based on the tensor representation of dFCNs. First, a framework for detecting change points at which the network connectivity patterns show consistent change across subjects is introduced through low-rank approximations to the tensors and computation of subspace distance between these low-rank approximations. Next, the network states corresponding to the different time intervals are extracted through tensor projection across subject and time modes. This approach is similar to PCA and extracting eigenvectors with significant energy. Finally, the proposed method is evaluated on EEG data collected during a study of cognitive control. The network states identified before, during, and after the ERN are consistent with hypothesized ERN-related network dynamics, focused on the formation of transient frontal networks during error processing followed by parietal networks.

The proposed approach offers several improvements over existing approaches. First, the proposed approach does not assume a discrete state model across time and extracts quasi-stationary network states similar to EEG microstates [80]. Second, by using the tensor representation of FCNs the topological structure of the networks are preserved and flexibility for identifying

network states across different variables such as frequency and stimulus type is introduced. Finally, it is shown that the proposed method is more sensitive to detecting change points and providing a more succinct summarization compared to matrix-based approaches such as SVD.

Future work will focus on several extensions. First, multiple tensor-matrix projections across subject and time modes will be considered and an optimal combination of these network states will be determined. This will allow each network state to have multiple components with different weights. Second, the network states from error and correct responses will be compared to better identify the role of control and salience networks.

APPENDIX A
PRINCIPAL ANGLE

Let $\mathbf{U}^{(1)}(t) = [\mathbf{u}_1^{(1)}(t)\ \mathbf{u}_2^{(1)}(t)\ \ldots\ \mathbf{u}_{\hat{n}(t)}^{(1)}(t)]$ and $\mathbf{U}^{(1)}(t + 1) = [\mathbf{u}_1^{(1)}(t+1)\ \mathbf{u}_2^{(1)}(t+1)\ \ldots\ \mathbf{u}_{\hat{n}(t+1)}^{(1)}(t+1)]$ span the subspaces of $\mathcal{G}(t)$ and $\mathcal{G}(t+1)$ along the first mode. Also, define $r = \min(\hat{n}(t), \hat{n}(t+1))$. Then, the kth principle vectors $\mathbf{p}_k \in \mathbf{U}^{(1)}(t)$ and $\mathbf{q}_k \in \mathbf{U}^{(1)}(t+1)$ for $k = 1, 2, \ldots, r$ are defined as a solution to the following optimization problem:

$$\mathrm{maximize}\ \mathbf{p}_k^\dagger \mathbf{q}_k,$$

$$\mathrm{subject\ to}$$

$$\mathbf{p}_k \in \mathbf{U}^{(1)}(t), \mathbf{q}_k \in \mathbf{U}^{(1)}(t+1),$$

$$\mathbf{p}_k^\dagger \mathbf{u}_1^{(1)}(t) = \cdots = \mathbf{p}_k^\dagger \mathbf{u}_{r-1}^{(1)}(t) = 0, \|\mathbf{p}_k\| = 1,$$

$$\mathbf{q}_k^\dagger \mathbf{u}_1^{(1)}(t+1) = \cdots = \mathbf{q}_k^\dagger \mathbf{u}_{r-1}^{(1)}(t+1) = 0, \|\mathbf{q}_k\| = 1.$$

$$(11)$$

The kth principal angle is defined as $\cos(\theta_k) = p_k^\dagger q_k; k = 1, 2, \ldots, r$[67].

REFERENCES

[1] M. Steriade *et al.*, "Basic mechanisms of cerebral rhythmic activities," *Electroencephalography Clinical Neurophysiol.*, vol. 76, no. 6, pp. 481–508, 1990.

[2] S. Bressler, "Large-scale cortical networks and cognition," *Brain Res. Rev.*, vol. 20, no. 3, pp. 288–304, 1995.

[3] F. Varela *et al.*, "The brainweb: Phase synchronization and large-scale integration," *Nature Rev. Neurosci.*, vol. 2, no. 4, pp. 229–239, 2001.

[4] F. Babiloni *et al.*, "Estimation of the cortical functional connectivity with the multimodal integration of high-resolution EEG and fMRI data by directed transfer function," *Neuroimage*, vol. 24, no. 1, pp. 118–131, 2005.

[5] A. Di Martino *et al.*, "Functional connectivity of human striatum: A resting state fMRI study," *Cerebral Cortex*, vol. 18, no. 12, pp. 2735–2747, 2008.

[6] M. Van Den Heuvel and H. Pol, "Exploring the brain network: A review on resting-state fMRI functional connectivity," *Eur. Neuropsychopharmacol.*, vol. 20, no. 8, pp. 519–534, 2010.

[7] E. Bullmore and O. Sporns, "Complex brain networks: Graph theoretical analysis of structural and functional systems," *Nature Rev. Neurosci.*, vol. 10, no. 3, pp. 186–198, 2009.

[8] M. van den Heuvel *et al.*, "Small-world and scale-free organization of voxel-based resting-state functional connectivity in the human brain," *Neuroimage*, vol. 43, no. 3, pp. 528–539, 2008.

[9] C. Stam and J. Reijneveld, "Graph theoretical analysis of networks in the brain," *Nonlinear Biomed. Phys.*, vol. 1, no. 1, p. 3, 2007.

[10] R. Prabhakaran *et al.*, "An event-related fMRI investigation of phonological–lexical competition," *Neuropsychologia*, vol. 44, no. 12, pp. 2209–2221, 2006.

[11] R. Goebel *et al.*, "Analysis of functional image analysis contest (FIAC) data with brainvoyager qx: From single-subject to cortically aligned group general linear model analysis and self-organizing group independent component analysis," *Human Brain Mapping*, vol. 27, no. 5, pp. 392–401, 2006.

[12] E. Bullmore and D. Bassett, "Brain graphs: Graphical models of the human brain connectome," *Annu. Rev. Clin. Psychol.*, vol. 7, pp. 113–140, 2011.

[13] P. Boveroux *et al.*, "Breakdown of within-and between-network resting state functional magnetic resonance imaging connectivity during propofol-induced loss of consciousness." *Anesthesiology*, vol. 113, no. 5, pp. 1038–53, 2010.

[14] J. Schrouff *et al.*, "Brain functional integration decreases during propofol-induced loss of consciousness," *Neuroimage*, vol. 57, no. 1, pp. 198–205, 2011.

[15] F. Freyer *et al.*, "Bistability and non-Gaussian fluctuations in spontaneous cortical activity," *J. Neurosci.*, vol. 29, no. 26, pp. 8512–8524, 2009.

[16] C. Chang and G. Glover, "Time–frequency dynamics of resting-state brain connectivity measured with fMRI," *Neuroimage*, vol. 50, no. 1, pp. 81–98, 2010.

[17] P. Kalivas and N. Volkow, "The neural basis of addiction: A pathology of motivation and choice," *Am. J. Psychiatry*, vol. 162, no. 8, pp. 1403–1413, 2005.

[18] K. Berridge and T. Robinson, "What is the role of dopamine in reward: Hedonic impact, reward learning, or incentive salience?" *Brain Res. Rev.*, vol. 28, no. 3, pp. 309–369, 1998.

[19] W. Seeley *et al.*, "Dissociable intrinsic connectivity networks for salience processing and executive control," *J. Neurosci.*, vol. 27, no. 9, pp. 2349–2356, 2007.

[20] E. Miller and J. Cohen, "An integrative theory of prefrontal cortex function," *Annu. Rev. Neurosci.*, vol. 24, no. 1, pp. 167–202, 2001.

[21] E. Miller, "The prefontral cortex and cognitive control," *Nature Rev. Neurosci.*, vol. 1, no. 1, pp. 59–65, 2000.

[22] R. Goldstein and N. Volkow, "Drug addiction and its underlying neurobiological basis: Neuroimaging evidence for the involvement of the frontal cortex," *Amer. J. Psychiatry*, vol. 159, no. 10, pp. 1642–1652, 2002.

[23] A. Bechara, "Decision making, impulse control and loss of willpower to resist drugs: A neurocognitive perspective," *Nature Neurosci.*, vol. 8, no. 11, pp. 1458–1463, 2005.

[24] R. Hutchison *et al.*, "Dynamic functional connectivity: Promise, issues, and interpretations," *Neuroimage*, vol. 80, pp. 360–378, 2013.

[25] J. Tang *et al.*, "Small-world behavior in time-varying graphs," *Phys. Rev. E*, vol. 81, no. 5, 2010. Art. no. 055101.

[26] P. Grindrod *et al.*, "Communicability across evolving networks," *Phys. Rev. E*, vol. 83, no. 4, 2011. Art. no. 046120.

[27] F. Fallani *et al.*, "Persistent patterns of interconnection in time-varying cortical networks estimated from high-resolution eeg recordings in humans during a simple motor act," *J. Phys. A, Math. Theoretical*, vol. 41, no. 22, 2008. Art. no. 224014.

[28] M. Valencia *et al.*, "Dynamic small-world behavior in functional brain networks unveiled by an event-related networks approach," *Phys. Rev. E*, vol. 77, no. 5, 2008. Art. no. 050905.

[29] S. Dimitriadis *et al.*, "Tracking brain dynamics via time-dependent network analysis," *J. Neurosci. Methods*, vol. 193, no. 1, pp. 145–155, 2010.

[30] S. Chiang *et al.*, "Time-dependence of graph theory metrics in functional connectivity analysis," *NeuroImage*, vol. 125, pp. 601–615, 2016.

[31] D. Bassett *et al.*, "Dynamic reconfiguration of human brain networks during learning," *Proc. Natl. Academy Sci. USA*, vol. 108, no. 18, pp. 7641–7646, 2011.

[32] E. Allen *et al.*, "Tracking whole-brain connectivity dynamics in the resting state," *Cerebral Cortex*, 2012. Art. no. bhs352.

[33] S. Dimitriadis *et al.*, "Transition dynamics of EEG-based network microstates during mental arithmetic and resting wakefulness reflects task-related modulations and developmental changes," *Cognitive Neurodynamics*, vol. 9, no. 4, pp. 371–387, 2015.

[34] S. Dimitriadis *et al.*, "On the quantization of time-varying phase synchrony patterns into distinct functional connectivity microstates (fcμstates) in a multi-trial visual ERP paradigm," *Brain Topography*, vol. 26, no. 3, pp. 397–409, 2013.

[35] W. Jamal *et al.*, "Existence of millisecond-order stable states in time-varying phase synchronization measure in EEG signals," in *Proc. IEEE 35th Annu. Int. Conf. Eng. Med. Biol. Soc.*, 2013, pp. 2539–2542.

[36] J. Ou *et al.*, "Characterizing and differentiating brain state dynamics via hidden Markov models," *Brain Topography*, vol. 28, no. 5, pp. 666–679, 2015.

[37] W. Jamal *et al.*, "Classification of autism spectrum disorder using supervised learning of brain connectivity measures extracted from synchrostates," *J. Neural Eng.*, vol. 11, no. 4, 2014. Art. no. 046019.

[38] W. Jamal *et al.*, "Using brain connectivity measure of EEG synchrostates for discriminating typical and autism spectrum disorder," in *Proc. 6th Int. IEEE/EMBS Conf. Neural Eng.*, 2013, pp. 1402–1405.

[39] W. Jamal *et al.*, "On the existence of synchrostates in multichannel eeg signals during face-perception tasks," *Biomed. Phys. Eng. Exp.*, vol. 1, no. 1, 2015. Art. no. 015002.

[40] S. Ma *et al.*, "Dynamic changes of spatial functional network connectivity in healthy individuals and schizophrenia patients using independent vector analysis," *NeuroImage*, vol. 90, pp. 196–206, 2014.

[41] R. Betzel *et al.*, "Synchronization dynamics and evidence for a repertoire of network states in resting EEG," *Frontiers iComputat. Neuroscience*, vol. 6, 2012.

[42] N. Leonardi *et al.*, "Disentangling dynamic networks: Separated and joint expressions of functional connectivity patterns in time," *Human Brain Mapping*, vol. 35, no. 12, pp. 5984–5995, 2014.

[43] N. Leonardi *et al.*, "Principal components of functional connectivity: A new approach to study dynamic brain connectivity during rest," *NeuroImage*, vol. 83, pp. 937–950, 2013.

[44] S. Mehrkanoon *et al.*, "Low-dimensional dynamics of resting-state cortical activity," *Brain Topography*, vol. 27, no. 3, pp. 338–352, 2014.

[45] A. Y. M. *et al.*, "A signal-processing-based approach to time-varying graph analysis for dynamic brain network identification," *Comput. Math. Methods Med.*, vol. 2012, p. 10, 2012.

[46] M. Lowe *et al.*, "Resting state sensorimotor functional connectivity in multiple sclerosis inversely correlates with transcallosal motor pathway transverse diffusivity," *Human Brain Mapping*, vol. 29, no. 7, pp. 818–827, 2008.

[47] I. Cribben *et al.*, "Dynamic connectivity regression: determining state-related changes in brain connectivity," *Neuroimage*, vol. 61, no. 4, pp. 907–920, 2012.

[48] I. Jolliffe, *Principal Component Analysis*. Hoboken, NJ, USA: Wiley, 2005.

[49] P. Comon, "Independent component analysis, a new concept?" *Signal Process.*, vol. 36, no. 3, pp. 287–314, 1994.

[50] E. Acar and B. Yener, "Unsupervised multiway data analysis: A literature survey," *IEEE Trans. Knowl. Data Eng.*, vol. 21, no. 1, pp. 6–20, Jan. 2009.

[51] L. De Lathauwer *et al.*, "A multilinear singular value decomposition," *SIAM J. Matrix Anal. Appl.*, vol. 21, no. 4, pp. 1253–1278, 2000.

[52] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, 2009.

[53] D. Goldfarb and Z. Qin, "Robust low-rank tensor recovery: Models and algorithms," *SIAM J. Matrix Anal. Appl.*, vol. 35, no. 1, pp. 225–253, 2014.

[54] A. Smilde *et al.*, *Multi-Way Analysis: Applications in the Chemical Sciences*. Hoboken, NJ, USA: Wiley, 2005.

[55] L. De Lathauwer, "A link between the canonical decomposition in multilinear algebra and simultaneous matrix diagonalization," *SIAM J. Matrix Anal. Appl.*, vol. 28, no. 3, pp. 642–666, 2006.

[56] L. De Lathauwer *et al.*, "Computation of the canonical decomposition by means of a simultaneous generalized Schur decomposition," *SIAM J. Matrix Anal. Appl.*, vol. 26, no. 2, pp. 295–327, 2004.

[57] H. Lu *et al.*, "A survey of multilinear subspace learning for tensor data," *Pattern Recog.*, vol. 44, no. 7, pp. 1540–1551, 2011.

[58] J. Kayser and C. Tenke, "Principal components analysis of Laplacian waveforms as a generic method for identifying ERP generator patterns: I. Evaluation with auditory oddball tasks," *Clinical Neurophysiol.*, vol. 117, no. 2, pp. 348–368, 2006.

[59] C. Tenke and J. Kayser, "Generator localization by current source density (CSD): Implications of volume conduction and field closure at intracranial and scalp resolutions," *Clinical Neurophysiol.*, vol. 123, no. 12, pp. 2328–2345, 2012.

[60] S. Aviyente and A. Mutlu, "A time-frequency-based approach to phase and phase synchrony estimation," *IEEE Trans. Signal Process.*, vol. 59, no. 7, pp. 3086–3098, Jul. 2011.

[61] M. Amft *et al.*, "Definition and characterization of an extended social-affective default network," *Brain Structure Function*, vol. 220, no. 2, pp. 1–19, 2015.

[62] D. Zoltowski and S. Aviyente, "Low-rank tensor decomposition based dynamic network tracking," in *Proc. IEEE Global Conf. Signal Inform. Process.*, 2014, pp. 468–472.

[63] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," in *Proc. Int. Joint Conf. Artif. Intell.*, vol. 14, no. 2, 1995, pp. 1137–1145.

[64] E. Ceulemans and H. Kiers, "Selecting among three-mode principal component models of different types and complexities: A numerical convex hull based method," *Brit. J. Math. Statist. Psychol.*, vol. 59, no. 1, pp. 133–150, 2006.

[65] L. Akoglu and C. Faloutsos, "Event detection in time series of mobile communication graphs," in *Proc. Army Sci. Conf.*, 2010, pp. 77–79.

[66] T. Ide and H. Kashima, "Eigenspace-based anomaly detection in computer systems," in *Proc. 10th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2004, pp. 440–449.

[67] K. Ye and L. Lim, "Distance between subspaces of different dimensions," arXiv:1407.0900, 2014.

[68] A. G. Mahyari and S. Aviyente, "Two-dimensional SVD for event detection in dynamic functional brain networks," in *Proc. IEEE Global Conf. Signal Inform. Process.*, 2013, pp. 37–40.

[69] M. G. Thomason and J. Gregor, "Higher order singular value decomposition of tensors for fusion of registered images," *J. Electron. Imag.*, vol. 20, no. 1, pp. 013 023–013 023, 2011.

[70] Q. Zhang *et al.*, "Multisensor video fusion based on higher order singular value decomposition," *J. Pers. Ubiquitous Comput.*, vol. 24, pp. 54–71, 2014.

[71] H. Lilliefors, "On the Kolmogorov–Smirnov test for normality with mean and variance unknown," *J. Amer. Statist. Assoc.*, vol. 62, no. 318, pp. 399–402, 1967.

[72] F. Massey Jr., "The Kolmogorov–Smirnov test for goodness of fit," *J. Amer. Statist. Assoc.*, vol. 46, no. 253, pp. 68–78, 1951.

[73] A. Karami *et al.*, "Compression of hyperspectral images using discrete wavelet transform and tucker decomposition," *IEEE J. Sel. Topics Appl. Earth Observations Remote Sens.*, vol. 5, no. 2, pp. 444–450, 2012.

[74] D. Donoho *et al.*, "Sparse solution of underdetermined systems of linear equations by stagewise orthogonal matching pursuit," *IEEE Trans. Inform. Theory*, vol. 58, no. 2, pp. 1094–1121, Feb. 2012.

[75] D. C. *et al.*, "R-MAT: A recursive model for graph mining," in *Proc. SIAM Int. Conf. Data Mining*, 2004, pp. 442–446.

[76] J. R. Hall, E. M. Bernat, and C. J. Patrick, "Externalizing psychopathology and the error-related negativity," *Psychological Sci.*, vol. 18, no. 4, pp. 326–333, 2007.

[77] J. Cavanagh *et al.*, "Prelude to and resolution of an error: EEG phase synchrony reveals cognitive control dynamics during action monitoring," *J. Neurosci.*, vol. 29, no. 1, pp. 98–105, 2009.

[78] S. Aviyente *et al.*, "A phase synchrony measure for quantifying dynamic functional integration in the brain," *Human Brain Mapping*, vol. 32, no. 1, pp. 80–93, 2011.

[79] M. Bolanos *et al.*, "A weighted small world network measure for assessing functional connectivity," *J. Neurosci. Methods*, vol. 212, no. 1, pp. 133–142, 2013.

[80] T. Koenig *et al.*, "Millisecond by millisecond, year by year: Normative EEG microstates and developmental stages," *Neuroimage*, vol. 16, no. 1, pp. 41–48, 2002.

Authors' photographs and biographies not available at the time of publication.