

Low-Rank Tensor Decomposition Based Dynamic Network Tracking

David M. Zoltowski, Selin Aviyente
Department of Electrical and Computer Engineering
Michigan State University
East Lansing, MI 48824, USA
zoltow11@msu.edu, aviyente@egr.msu.edu

Abstract—Lots of data is generated around us in today’s big data age. Much of this data is time-varying or dynamic, such as the social network connections that change across time or dynamic functional brain connectivity networks constructed across multiple subjects. In these dynamic (time-varying) networks, it is important to reduce the large amount of data into a few meaningful descriptors. One way to achieve this goal is to detect change points or anomalies in the connectivity patterns across time. Recently, there has been an interest in describing the time-varying network activity as a tensor and detecting the anomalies in terms of the changes in the subspaces of the tensor along each mode [1], [2]. However, the current approaches to tensor decomposition are not robust to non-Gaussian noise, outliers, and corruption in the data. For this reason, a robust low-rank tensor recovery algorithm similar to robust principal components analysis (RPCA) has been recently proposed. In this paper, we employ higher order robust PCA (HoRPCA) for tracking dynamic networks in time and detecting anomalies using a subspace distance measure. The proposed approach assumes that most real life networks are low-rank in nature and considers a low-rank plus sparse tensor decomposition at each time point. The subspaces corresponding to each mode and each time point are described through a projection operator and the subspace distance is quantified through a Hausdorff distance measure. The proposed framework is evaluated on both simulated networks and dynamic functional connectivity brain networks.

Index Terms—dynamic networks, tensor decomposition, robust PCA, functional connectivity

I. INTRODUCTION

Rapid advances in information technology have resulted in an abundance of data. Much of this data is multidimensional or tensor data. Tensors are the higher-order generalization of vectors and matrices and have found many applications in the physical and imaging sciences, web data mining, neuroscience and signal processing [3], [4]. Traditional matrix-based data analysis is inherently two-dimensional, which limits its usefulness in extracting information from a multidimensional perspective. On the other hand, tensor-based multilinear data analysis has shown that tensor models are capable of taking full advantage of the multilinear structures by extracting hidden structures, capturing underlying correlations between variables and providing higher precision approximations.

Although these multilinear models have been extensively used in modeling or describing a variety of higher order data [5], [6], [7], [8], their primary use has been to reduce large amounts of data across different parameters into a small num-

ber of components. Recently, there has also been a growing interest in efficient implementation of these models for big data [1], [2], [9]. However, in most of these applications, it is assumed that the whole dataset is available and the analysis is done offline. In the area of dynamic networks it is important to track the changes to the network structure as new data becomes available because these changes may correspond to anomalies in the network. Sun et al. [10] proposed using tensors for tracking dynamic networks, where they computed the projection matrices along each mode recursively and used reduced-rank projection matrices for approximating the tensor and the residual error for detecting anomalies. More recently, non-negative Tucker decomposition has been used to identify communities in social networks across time [11].

In this paper, we extend the previous approaches in some key ways. First, unlike current approaches which are mostly limited to time-varying social networks, thus 3-way tensors, we consider higher order tensors where other experimental parameters such as the different modalities that the same group of people interact over can be taken into account. Second, most of the current network tracking approaches rely on Tucker decomposition and the low rank approximations to the data for detecting anomalies. However, when the data is corrupted by non-Gaussian noise or when there are outliers in the data, Tucker decomposition cannot reveal the actual low-rank structure of the data. Moreover, it is known that although lower rank truncated representations are computationally efficient, they do not necessarily provide the optimal approximation. For this reason, in this paper we utilize a recently introduced higher order robust PCA (HoRPCA) approach for recovering the low-rank part from the observed tensor [12]. Finally, instead of tracking based on residual errors, which is optimal in the case of Gaussian noise, we propose to use a subspace distance based approach to network tracking.

In this paper, we focus on the application of this approach to dynamic functional connectivity networks. A key advance within neuroscience is mapping the functional connectivity (FC) networks within the brain [13], [14]. Functional connectivity can be defined as the statistical dependency between spatially remote brain regions, and characterizes interregional neural interactions known to be responsible for cognitive processes [15], [16], [17]. Current functional connectivity studies assume the temporal stationarity of the network during the

recording period, thus representing an average across complex spatio-temporal phenomena. A growing number of studies now make clear that fluctuation of FC due to task demands, learning, sleep, and anesthesia, as well as during the resting state, are critical for accurately characterizing these activities [18], [19].

II. BACKGROUND

A. Tensor Decomposition

At the core of multilinear data analysis lies tensor decomposition, which commonly takes two forms: CANDECOMP/PARAFAC (CP) decomposition and Tucker decomposition [4]. The goal of tensor decomposition in data reduction applications is to find a good multilinear rank (r_1, r_2, \dots, r_d) approximation to the order- d tensor $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_d}$, i.e. $\min_{\mathcal{B}} \|\mathcal{A} - \mathcal{B}\|_F^2$ with $\mathcal{B} \in \mathbb{R}^{n_1 \times \dots \times n_d}$ restricted to be of rank (r_1, r_2, \dots, r_d) and $r_i \leq n_i$. This approximation problem is well-posed [20], but does not have a known closed solution. Different algorithms have been proposed as a solution to this problem, the most well-known being the higher-order singular value decomposition (HOSVD) [20], [21] which is practically useful for constructing an approximate solution to this problem. A reduced-rank approximation is simply obtained by restricting the factor matrices of \mathcal{A} , $\mathbf{U}_i \in \mathbb{R}^{n_i \times n_i}$, where $\mathcal{A} = \mathcal{C} \times_1 \mathbf{U}_1 \times_2 \mathbf{U}_2 \times_3 \dots \times_d \mathbf{U}_d$ and $\mathcal{C} \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_d}$ is the core tensor, to the first r_1, r_2, \dots, r_d columns.

B. Time-Varying Functional Networks

The time-varying functional brain networks are generated where the nodes of the graphs correspond to different brain regions and the edges correspond to the connectivity strengths. In this paper, we quantify the connectivity using a recent phase synchrony measure based on RID-Rihaczek distribution [22]. This measure has been shown to be more robust to noise and to provide better resolution as discussed in [22].

First, we quantify the time-varying phase of a signal, $\Phi_i(t, \omega) = \arg \left[\frac{C_i(t, \omega)}{|C_i(t, \omega)|} \right]$ where $C_i(t, \omega)$ is the complex RID-Rihaczek distribution¹

$$C_i(t, \omega) = \int \int \underbrace{\exp\left(-\frac{(\theta\tau)^2}{\sigma}\right)}_{\text{Choi-Williams kernel}} \underbrace{\exp\left(j\frac{\theta\tau}{2}\right)}_{\text{Rihaczek kernel}} A_i(\theta, \tau) e^{-j(\theta t + \tau\omega)} d\tau d\theta \quad (1)$$

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 s_i . The phase synchrony between nodes i and j at time t and frequency ω is computed using ‘Phase Locking Value’ (PLV):

$$PLV_{i,j}(t, \omega) = \frac{1}{L} \left| \sum_{k=1}^L \exp(j\Phi_{i,j}^k(t, \omega)) \right| \quad (2)$$

where L is the number of trials and $\Phi_{i,j}^k(t, \omega) = |\Phi_i(t, \omega) - \Phi_j(t, \omega)|$ is the phase difference estimate between the two nodes for the k^{th} trial.

¹The details of the RID-Rihaczek distribution and the corresponding synchrony measure are given in [22].

Let $\{\mathbf{G}(t)\}_{t=1,2,\dots,T}$ be a time sequence of graphs, where $\mathbf{G}(t)$ is an $N \times N$ weighted and undirected graph corresponding to the functional connectivity network at time t for a fixed frequency or frequency band, T is the total number of time points and N is the number of nodes within the network. The time-varying edge values are quantified by the average PLV within a frequency band and at a certain time as:

$$G_{i,j}(t) = \frac{1}{\Omega} \sum_{\omega=\omega_a}^{\omega_b} PLV_{i,j}(t, \omega) \quad (3)$$

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.

III. METHODS

A. Robust Low-Rank Tensor Recovery Algorithms

In practice, the underlying tensor data is often low-rank, even though the actual observed data may not be due to outliers and arbitrary errors. Therefore, the major part of the variation in the data is often governed by a relatively small number of latent factors. In order to address this problem, we will make the assumption that at each time point t the three-way tensor $\mathcal{G}_t \in \mathbb{R}^{N \times N \times S}$ is low-rank and is composed of a sparse tensor and a low-rank tensor. In this case, S is the number of subjects and $t = 1, 2, \dots, T$. We assume that $\mathcal{G}_t = \mathcal{S}_t + \mathcal{L}_t$, where \mathcal{S}_t is a sparse tensor and \mathcal{L}_t is a dense and low-rank tensor. Moreover, \mathcal{L}_t lives in a low-dimensional manifold that changes slowly with time, whereas \mathcal{S}_t corresponds to more transient and spontaneous activity. In the context of the brain connectivity \mathcal{L}_t would correspond to the resting-state network which can be assumed to be stationary or slowly changing with time whereas \mathcal{S}_t would correspond to the activated connections due to a stimulus.

For a fixed t , to separate the low-rank structure from noisy tensor data subject to outliers and arbitrary corruptions, we will employ an extension of robust PCA (RPCA) [23] to higher-order data models. A direct application of RPCA considers the low-rank structure in only one of the unfoldings and is often insufficient. To generalize RPCA to tensors, several extensions [12], [24] have been proposed by regularizing the Tucker-rank of a tensor leading to the following optimization problem

$$\begin{aligned} \min_{\mathcal{S}, \mathcal{L}} \quad & \text{Trank}(\mathcal{L}) + \lambda \|\mathcal{S}\|_0 \\ \text{such that} \quad & \mathcal{G} = \mathcal{S} + \mathcal{L}. \end{aligned} \quad (4)$$

This problem is NP-hard, so a convex version of this optimization problem is used in practice as

$$\min_{\mathcal{S}, \mathcal{L}} \left\{ \sum_{i=1}^3 \|L_{(i)}\|_* + \lambda_1 \|\mathcal{S}\|_1 \mid \mathcal{S} + \mathcal{L} = \mathcal{G} \right\}, \quad (5)$$

where L_i is the mode- i unfolding (matricization) of \mathcal{L} and $\|\cdot\|_*$ is the nuclear norm of the matrix. The nuclear norm is defined as $\sum_{i=1}^r \sigma_i$ where σ_i s are the singular values of the rank r matrix and the ℓ_1 norm of a tensor is defined as the ℓ_1 norm

of the vectorized form of the tensor. This problem has been previously solved using an alternating direction augmented Lagrangian (ADAL) method using variable-splitting and is also known as the Higher-order RPCA (HoRPCA) [12]. Once the optimal \mathcal{L} is obtained, we can reconstruct the core tensor \mathcal{C} of \mathcal{L} by $\mathcal{C} = \mathcal{L} \times_1 (\mathbf{U}^{(1)})^T \times_2 (\mathbf{U}^{(2)})^T \times_3 (\mathbf{U}^{(3)})^T$, where $\mathbf{U}^{(i)}$ is the left factor matrix from the SVD of $\mathbf{L}^{(i)}$. Thus, we can recover the Tucker decomposition, i.e. the subspaces along each mode, of the low-rank tensor \mathcal{L} from its corrupted version without the need to specify the target Tucker-rank.

In particular, introducing three auxiliary variables $\mathcal{L}_1, \mathcal{L}_2, \mathcal{L}_3$ to variable-split \mathcal{L} reformulates the convex optimization problem as

$$\min_{\mathcal{S}, \mathcal{L}_1, \mathcal{L}_2, \mathcal{L}_3} \left\{ \sum_{i=1}^3 \|\mathbf{L}_{i,(i)}\|_* + \lambda_1 \|\mathcal{S}\|_1 \mid \mathcal{S} + \mathcal{L}_i = \mathcal{G} \right\}.$$

The augmented Lagrangian for this problem is

$$\begin{aligned} & \sum_{i=1}^3 \|\mathbf{L}_{i,(i)}\|_* + \lambda_1 \|\mathcal{S}\|_1 + \\ & \sum_{i=1}^3 \left(\frac{1}{2\mu} \|\mathcal{L}_i + \mathcal{S} - \mathcal{G}\|^2 - \langle \Lambda_i, \mathcal{L}_i + \mathcal{S} - \mathcal{G} \rangle \right), \end{aligned} \quad (6)$$

where Λ_i is an update variable and the \mathcal{L}_i 's can be solved for independently by the singular value thresholding operator with threshold μ for a given \mathcal{S} . On the other hand, with fixed \mathcal{L}_i 's, the augmented Lagrangian subproblem with respect to \mathcal{S} has the closed-form solution

$$\mathcal{S} = \mathcal{T}_{\frac{\mu\lambda_1}{3}} \left(-\frac{1}{3} \sum_{i=1}^3 (\mathcal{L}_i - \mathcal{G} - \mu\Lambda_i) \right),$$

where the shrinkage operator \mathcal{T} soft-thresholds the elements of the tensor depending on the regularization parameters μ and λ_1 . In our implementation, we investigate the effect of these two parameters on the resulting low-rank tensor decomposition.

B. Low-Rank Tensor Decomposition Based Change Point Detection

In this paper, we extend the HoRPCA algorithm to the dynamic case to identify the change points. Change points will be identified through the change in the subspaces along each mode of \mathcal{L}_t . Assuming the low-rank tensor \mathcal{L}_t is slowly changing, at each time point we will measure the subspace distance along each mode of the tensor. We define the projection operator used in our subspace distance measure as $\mathbf{P}_i^{(t)} = \mathbf{U}_i^{(t)} \mathbf{U}_i^{(t)T}$ along mode i and at time t . Quantifying the subspace distance through the Hausdorff distance, which takes into account the difference in the dimensions between subspaces [25], the change between the subspaces of \mathcal{L}_t and \mathcal{L}_{t+1} will be defined as

$$\sum_{i=1}^3 \sqrt{\max(r_i^{(t)}, r_i^{(t+1)}) - \text{tr}(\mathbf{P}_i^{(t)} \mathbf{P}_i^{(t+1)})}, \quad (7)$$

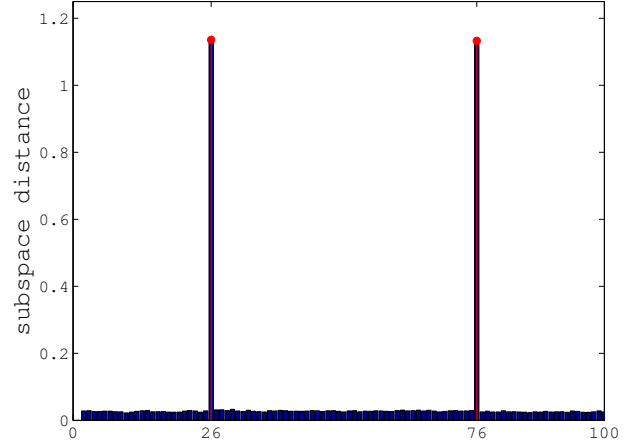


Fig. 1: The subspace distances between the projection operators at consecutive time points for $\alpha = 1/4$. Points in red denote change points that are greater than two standard deviations from the mean subspace distance.

where $r_i^{(t)}$ is the dimension of the subspace for the i th mode at time t . This total change between subspaces will be used to detect changes in time and segment the time duration into partitions over which the network is quasi-stationary. This tracking will be based on statistical significance or thresholding similar to [26], [27], [28]. Specifically, when a time point t has a Hausdorff distance (7) greater than the mean Hausdorff distance plus two standard deviations, t is selected as a change point.

IV. RESULTS

In this section, we will evaluate the performance of the proposed dynamic network tracking algorithm on simulated tensors and then on time-varying functional brain connectivity graphs.

A. Simulated Data

We generated a four-way tensor corresponding to a social network where the first two dimensions represented connections between $N = 32$ subjects, the third dimension represented $M = 50$ modalities, and the fourth dimension represented $T = 100$ time points. We designed the tensor such that it contained two change points at $T_1 = 26$ and $T_2 = 76$. In the interval $[0, 25]$, we created two clusters of subject connections, where we selected the within-cluster edge weights randomly from the uniform distribution $[0.8, 1]$ and the between-cluster edge weights randomly from the uniform distribution $[0, 0.2]$. In the interval $[26, 75]$, we designed the graphs to have three clusters instead of two clusters using the same within- and between-cluster edge weight distributions. Finally, for the interval $[76, 100]$, we reverted to the original graph structure with two clusters of subject connections.

For each time point t , we performed Higher-order Robust PCA on the resulting tensor $\mathcal{X}_t \in R^{32 \times 32 \times 50}$. We chose

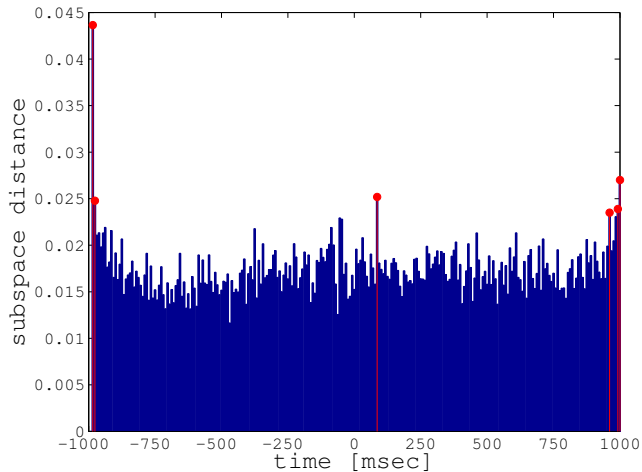


Fig. 2: The subspace distances between the projection operators at consecutive time points for $\alpha = 1/4$. Points in red denote change points that are greater than two standard deviations from the mean subspace distance.

the Singleton model, where the tensor rank regularization parameter is the sum of the nuclear norms [12], and set the algorithm parameters $\lambda_1 = \frac{\alpha \lambda_*}{\sqrt{I_{max}}}$ and $\mu = 10$. We chose $\alpha = \frac{1}{4}$, $\lambda_* = 1$, and set I_{max} equal to the maximum of the sizes of the dimensions of \mathcal{X}_t , in this case 50. The choice of parameters promoted low rank tensor recovery at the expense of sparsity. Subspace distances were calculated between the projection operators at consecutive time points and plotted in Fig. 1. The algorithm detected change points at $t = 26$ and $t = 76$, which corresponds exactly to the cluster structure change points for the simulated data.

B. EEG Data

The proposed framework is applied to a set of electroencephalogram (EEG) data containing the error-related negativity (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 [29]. Previous work [30] indicates that there is increased coordination between the lateral prefrontal cortex (IPFC) and medial prefrontal cortex (mPFC) within the theta frequency band (4-8 Hz) and ERN time window (25- 75 ms), supporting the idea that frontal and central electrodes are functionally integrated during error processing. EEG data from 63 channels with a sampling rate of 128 Hz was collected in accordance with the 10/20 system on a Neuroscan Synamps2 system (Neuroscan, Inc.). A speeded-response flanker task was employed, and response-locked averages were computed for each subject across 2 seconds. All EEG epochs were converted to current source density (CSD) using published methods [31]. In this paper, we analyzed data from 91 subjects corresponding to the error responses.

We constructed the time-varying edge values for our graphs

with $N = 63$, $T = 256$, and $S = 91$ by computing the average PLV within the frequency band [2, 7] Hz for each pair of nodes at a certain time and for a certain subject. For each time point t , we performed Higher-order Robust PCA on the resulting tensor $\mathcal{G}_t \in R^{63 \times 63 \times 91}$. We again used a Singleton model and set the tuning parameters $\alpha = \frac{1}{4}$, $\lambda_* = 1$, and $\mu = 2$.

Subspace distances were calculated between the projection operators at consecutive time points and are plotted across time in Fig. 2 with detected change points, or values that were larger than two standard deviations from the mean, denoted in red. Based on this analysis, we detected a change point at 85.9375 ms along with change points at the beginning and end time ranges of the recording period. It is important to note that the change point at 85.9375 ms closely corresponds to the ERN time window indicating a significant change in the network structure.

V. CONCLUSIONS

In this paper, we introduced a Higher Order Robust PCA (HoRPCA) based approach combined with a subspace distance measure to track the change points in dynamic higher order data sets. This type of data occurs commonly in social networks where the connectivity may be collected across different modalities between subjects across time resulting in a multilinear data set as well as in functional connectivity networks constructed from neuroimaging data. The proposed approach is based on solving a convex optimization problem to recover the low-rank tensor from noisy multilinear data. The regularization parameters in the optimization problem are selected to achieve a balance between the rank and the sparsity.

Future work will consider recursive implementation of HoRPCA in the case of slowly changing subspaces of \mathcal{L} increasing the efficiency of the approach. Moreover, network summarization based on the sparse component in the decomposition, \mathcal{S} , will be considered.

ACKNOWLEDGEMENT

This work was in part supported by the National Science Foundation under Grant No. CCF-1218377.

REFERENCES

- [1] A. Cichocki, "Era of big data processing: A new approach via tensor networks and tensor decompositions," *arXiv preprint arXiv:1403.2048*, 2014.
- [2] M. Mardani, G. Mateos, and G. B. Giannakis, "Subspace learning and imputation for streaming big data matrices and tensors," *arXiv preprint arXiv:1404.4667*, 2014.
- [3] E. Acar and B. Yener, "Unsupervised multiway data analysis: A literature survey," *Knowledge and Data Engineering, IEEE Transactions on*, vol. 21, no. 1, pp. 6–20, 2009.
- [4] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM review*, vol. 51, no. 3, pp. 455–500, 2009.
- [5] L. Omberg, G. H. Golub, and O. Alter, "A tensor higher-order singular value decomposition for integrative analysis of dna microarray data from different studies," *Proceedings of the National Academy of Sciences*, vol. 104, no. 47, pp. 18 371–18 376, 2007.
- [6] M. Mørup, L. K. Hansen, C. S. Herrmann, J. Parnas, and S. M. Arnfred, "Parallel factor analysis as an exploratory tool for wavelet transformed event-related EEG," *NeuroImage*, vol. 29, no. 3, pp. 938–947, 2006.
- [7] H. Lee, Y.-D. Kim, A. Cichocki, and S. Choi, "Nonnegative tensor factorization for continuous eeg classification," *International journal of neural systems*, vol. 17, no. 04, pp. 305–317, 2007.

- [8] F. Miwakeichi, E. Martinez-Montes, P. A. Valdés-Sosa, N. Nishiyama, H. Mizuhara, and Y. Yamaguchi, "Decomposing eeg data into space-time-frequency components using parallel factor analysis," *NeuroImage*, vol. 22, no. 3, pp. 1035–1045, 2004.
- [9] N. Sidiropoulos, E. Papalexakis, and C. Faloutsos, "A parallel algorithm for big tensor decomposition using randomly compressed cubes (para-comp)."
- [10] J. Sun, D. Tao, and C. Faloutsos, "Beyond streams and graphs: dynamic tensor analysis," in *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2006, pp. 374–383.
- [11] L. Gauvin, A. Panisson, and C. Cattuto, "Detecting the community structure and activity patterns of temporal networks: a non-negative tensor factorization approach," *PLoS one*, vol. 9, no. 1, p. e86028, 2014.
- [12] D. Goldfarb and Z. Qin, "Robust low-rank tensor recovery: Models and algorithms," *SIAM Journal on Matrix Analysis and Applications*, vol. 35, no. 1, pp. 225–253, 2014.
- [13] B. He, T. Coleman, G. M. Genin, G. Glover, X. Hu, N. Johnson, T. Liu, S. Makeig, P. Sajda, and K. Ye, "Grand challenges in mapping the human brain: NSF workshop report," *IEEE Transactions on Biomedical Engineering*, in press.
- [14] K. Friston, "Functional and effective connectivity in neuroimaging: a synthesis," *Human Brain Mapping*, vol. 2, no. 1-2, pp. 56–78, 1994.
- [15] P. J. Uhlhaas and W. Singer, "Neural synchrony in brain disorders: Relevance for cognitive dysfunctions and pathophysiology," *Neuron*, vol. 52, pp. 155–168, 2006.
- [16] K. M. Spencer, P. G. Nestor, M. A. Niznikiewicz, D. F. Salisbury, M. E. Shenton, and R. W. McCarley, "Abnormal neural synchrony in schizophrenia," *The Journal of Neuroscience*, vol. 23, no. 19, pp. 7407–7411, 2003.
- [17] V. Menon, "Large-scale brain networks and psychopathology: a unifying triple network model," *Trends in cognitive sciences*, vol. 15, no. 10, pp. 483–506, 2011.
- [18] C. Chang and G. H. Glover, "Time-frequency dynamics of resting-state brain connectivity measured with fmri," *Neuroimage*, vol. 50, no. 1, pp. 81–98, 2010.
- [19] N. Leonardi, J. Richiardi, M. Gschwind, S. Simioni, J.-M. Annoni, M. Schluep, P. Vuilleumier, and D. Van De Ville, "Principal components of functional connectivity: A new approach to study dynamic brain connectivity during rest," *NeuroImage*, vol. 83, pp. 937–950, 2013.
- [20] L. De Lathauwer, B. De Moor, and J. Vandewalle, "On the best rank-1 and rank-(r_1, r_2, \dots, r_n) approximation of higher-order tensors," *SIAM Journal on Matrix Analysis and Applications*, vol. 21, no. 4, pp. 1324–1342, 2000.
- [21] —, "A multilinear singular value decomposition," *SIAM journal on Matrix Analysis and Applications*, vol. 21, no. 4, pp. 1253–1278, 2000.
- [22] S. Aviyente and A. Mutlu, "A time-frequency based approach to phase and phase synchrony estimation," *Signal Processing, IEEE Transactions on*, vol. 59, no. 7, pp. 3086–3098, 2011.
- [23] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *Journal of the ACM (JACM)*, vol. 58, no. 3, p. 11, 2011.
- [24] S. Gandy, B. Recht, and I. Yamada, "Tensor completion and low-n-rank tensor recovery via convex optimization," *Inverse Problems*, vol. 27, no. 2, p. 025010, 2011.
- [25] G. Zuccon, L. A. Azzopardi, and C. Van Rijsbergen, "Semantic spaces: Measuring the distance between different subspaces," in *Quantum Interaction*. Springer, 2009, pp. 225–236.
- [26] A. Lakhina, M. Crovella, and C. Diot, "Diagnosing network-wide traffic anomalies," in *ACM SIGCOMM Computer Communication Review*, vol. 34, no. 4. ACM, 2004, pp. 219–230.
- [27] —, "Mining anomalies using traffic feature distributions," in *ACM SIGCOMM Computer Communication Review*, vol. 35, no. 4. ACM, 2005, pp. 217–228.
- [28] L. Akoglu and C. Faloutsos, "Event detection in time series of mobile communication graphs," in *Army Science Conference*, 2010.
- [29] J. R. Hall, E. M. Bernat, and C. J. Patrick, "Externalizing psychopathology and the error-related negativity," *Psychological Science*, vol. 18, no. 4, pp. 326–333, 2007.
- [30] J. Cavanagh, M. Cohen, and J. Allen, "Prelude to and resolution of an error: Eeg phase synchrony reveals cognitive control dynamics during action monitoring," *The Journal of Neuroscience*, vol. 29, no. 1, pp. 98–105, 2009.
- [31] 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 Neurophysiology*, vol. 117, no. 2, pp. 348–368, 2006.