Learning Spatiotemporal Brain Dynamics in Adolescents via Multimodal MEG and fMRI Data Fusion Using Joint Tensor/Matrix Decomposition

Irina Belyaeva, *Member, IEEE*, Ben Gabrielson, *Graduate Student Member, IEEE*, Yu-Ping Wang, *Senior Member, IEEE*, Tony W. Wilson, Vince D. Calhoun, *Fellow, IEEE*, Julia M. Stephen, and Tülay Adali, *Fellow, IEEE*

***Abstract*— Objective: Brain function is understood to be regulated by complex spatiotemporal dynamics, and can be characterized by a combination of observed brain response patterns in time and space. Magnetoencephalography (MEG), with its high temporal resolution, and functional magnetic resonance imaging (fMRI), with its high spatial resolution, are complementary imaging techniques with great potential to reveal information about spatiotemporal brain dynamics. Hence, the complementary nature of these imaging techniques holds much promise to study brain function in time and space, especially when the two data types are allowed to fully interact. *Methods:* We employed coupled tensor/matrix factorization (CMTF) to extract joint latent components in the form of unique spatiotemporal brain patterns that can be used to study brain development and function on a millisecond scale. *Results:* Using the CMTF model, we extracted distinct brain patterns that revealed fine-grained spatiotemporal brain dynamics and typical sensory processing pathways informative of high-level cognitive functions in healthy adolescents. The components extracted from multimodal tensor fusion possessed better discriminative ability between high- and low-performance subjects than single-modality data-driven models. *Conclusion:* Multimodal tensor fusion successfully identified spatiotemporal brain dynamics of brain function and produced unique components with high discriminatory power. *Significance:* The CMTF model is a promising tool for high-order, multimodal data fusion that exploits the functional resolution of MEG and fMRI, and provides a comprehensive picture of the developing brain in time and space.**

***Index Terms*—Brain function, coupled tensor/matrix factorization, developmental neuroscience, fMRI, MEG, multimodal data fusion, spatiotemporal brain dynamics.**

I. Introduction

MULTIMODAL neuroimaging is an important tool for neuroscience research, due to the complex spatiotemporal dynamics being characterized by the collective information present across multiple imaging modalities. Multimodal data fusion can take advantage of multiple neuroimaging techniques to improve the characterization of the temporal and spatial resolution of neurological processes, and identify pathology with high sensitivity and specificity [1], [2]. Magnetoencephalography (MEG) and electroencephalography (EEG) are direct neuroimaging techniques that instantaneously measure the neural currents induced by neuronal activity. MEG is a powerful neuroimaging technique for studying the rich temporal dynamics of neuronal activity with submillisecond resolution. Functional magnetic resonance imaging (fMRI) is an indirect neuroimaging technique that measures local changes in brain hemodynamics with a high spatial resolution. However, the temporal resolution of fMRI is limited to a few hundred milliseconds. Given noticeable complementarity between the two techniques, integrating MEG and fMRI has become a highly desirable multimodal approach [3]. Therefore, joint analysis of MEG and fMRI is of the greatest interest for enabling a comprehensive large-scale view of brain function. State-of-the-art methods for neuroimaging data joint analysis comprise different approaches to data fusion, or approaches that estimate the interaction between different modalities [4]. Several data fusion methods have been proposed for fusion of the MEG/EEG multimodal data based on matrix and tensor decompositions. These include joint independent component analysis (jICA) [5], multiset canonical correlation analysis mCCA [6], dictionary learning (DL) [7], and independent vector analysis (IVA) and its transposed variant (tIVA) [8] among others [9]. Data-driven matrix decomposition methods are powerful because they can extract meaningful biological

Manuscript received 13 December 2022; revised 9 August 2023, 6 November 2023, and 16 January 2024; accepted 6 February 2024. Date of publication 12 February 2024; date of current version 20 June 2024. This work was supported in part by NSF-NCS under Grant 1631838, in part by NSF under Grant 2112455, and in part by NIH under Grant R01 MH118695, Grant R01MH123610, Grant R01AG073949, Grant R01 MH121101, and Grant P20 GM144641. *(Corresponding author: Irina Belyaeva.)*

Irina Belyaeva is with the University of Maryland, Baltimore County, Baltimore, MD 21250 USA (e-mail: irinbel1@umbc.edu).

Ben Gabrielson and Tülay Adali are with the Department of Computer Science and Electrical Engineering, University of Maryland, Baltimore County, USA.

Yu-Ping Wang is with the Department of Biomedical Engineering, Tulane University, USA.

Tony W. Wilson is with the Institute for Human Neuroscience, Boys Town National Research Hospital, USA.

Vince D. Calhoun is with the Tri-institutional Center for Translational Research in Neuroimaging and Data Science, USA.

Julia M. Stephen is with the Mind Research Network a Division of Lovelace Biomedical Research Institute, USA.

This article has supplementary downloadable material available at https://doi.org/10.1109/TBME.2024.3364704, provided by the authors.

Digital Object Identifier 10.1109/TBME.2024.3364704

patterns across multiple modalities [10] and enable fusion using second order or high-order statistics. However, there are two drawbacks of matrix decomposition methods: (1) flattening naturally high-dimensional data into a two-dimensional (2D) representation does not fully leverage the multilinear structure of multisensor MEG/EEG, and (2) enforcing constraints such as independence or sparsity of components is typically required to ensure the uniqueness of the matrix decomposition [1], [7]. In contrast, multiway data fusion methods address the drawbacks of matrix decomposition methods by jointly analyzing datasets in the form of matrices and tensors without imposing additional constraints, and account for the multilinear structure of naturally high-order MEG/EEG. The multiway decomposition methods for MEG/EEG and fMRI data fusion include N-way partial least squares [11], coupled matrix/tensor factorization (CMTF) [12] and coupled tensor-tensor factorization (CTTD) [13]. Technically, both CMTF and CTTD decomposition methods employ canonical polyadic (CP) decomposition [14], which is essentially unique up to scaling and permutations [15].

Fusion of EEG and fMRI data has demonstrated the ability to address the inherent limitations of each imaging modality, resulting in enhanced spatiotemporal resolution and improved localization of neural sources [16], [17], [18]. Furthermore, EEG and fMRI data fusion offer significant benefits, such as high sensitivity to radially-oriented sources, the ability to detect deep sources in medial and subcortical regions, and the availability of the method in hospitals when translating studies into clinical practice [19], [20]. Despite these advantages, the fusion of EEG and fMRI data still faces significant limitations. A notable challenge is the presence of dominant artifacts, including gradient artifact and ballistocardiogram (BCG), that arise during the simultaneous acquisition of these modalities [21], [22]. Another limitation stems from the intrinsic issues associated with EEG recordings, particularly volume conduction. Volume conduction introduces complexities in relating EEG data to other imaging modalities. The distortion of EEG signals caused by a poorly conducting skull poses challenges to accurately identify the underlying neural sources, leading to errors in source localization [23], [24], [25], [26]. Unlike EEG, magnetic fields propagate through the head almost unchanged and thus experience minimal spatial distortion. For this reason, MEG recordings result in significantly lower source localization errors [23], [25], [27], offer a higher spatial resolution than EEG, and thus MEG and fMRI data fusion can separate neural sources with a higher specificity [20]. Furthermore, since MEG is biased towards tangential sources, it provides increased sensitivity to the detection of activity arising from the fissural cortex [20]. Therefore, the selectivity of MEG and fMRI data fusion to tangential sources in the presence of several simultaneous sources is a significant advantage in differentiating between multiple cortical areas activated by somatosensory stimuli [20]. However, MEG demonstrates a decrease in sensitivity to deeper sources [20], [25]. Although EEG is not affected by the lower sensitivity to deep sources as MEG and offers a more isotropic sensitivity profile, the spread of electrical potential on the scalp hinders source localization [28]. That said, both MEG and EEG provide high temporal resolution (on the order of milliseconds), which is complementary to the hemodynamic response measured with fMRI (on the order of seconds). Because of the different time scales of electrophysiological (MEG/EEG) vs. hemodynamic (fMRI) techniques, we expect each technique will capture both unique and overlapping spatiotemporal information. For example, with high temporal resolution MEG can capture a sequence of events within the brain by examining different time points while performing tasks. However, deep sources produce weaker signals at the MEG sensors limiting sensitivity to these structures, whereas fMRI provides sensitivity to both cortical and subcortical structures but cannot separately capture the sequence of events at the millisecond timescale. Furthermore, the combination of MEG and fMRI holds promising prospects in noise reduction inherent to each modality [22], [28], [29], [30]. As a result, MEG and fMRI data fusion shows great potential for substantially boosting statistical sensitivity and outperforming the use of either modality alone. This promise arises from several key factors: MEG's superior source localization accuracy over EEG, the improved spatiotemporal resolution achieved through the fusion of MEG and fMRI data, and the noise reduction capabilities of the fusion process. Consequently, this integrated approach enables more effective discrimination between experimental conditions, making it a powerful tool in neuroscientific research.

Given the advantages of MEG over EEG and the complementarity of MEG and fMRI, and the inherent multiway nature of MEG data, tensor-based fusion of these modalities is particularly attractive. The joint analysis of MEG and fMRI provides a comprehensive view of brain processes beyond what each imaging method can achieve independently. Therefore, the fusion of MEG and fMRI holds great potential to significantly contribute to the current state-of-the-art multimodal imaging methods and further deepen our understanding of the human brain and cognitive function.

The analysis of MEG/EEG data using tensor factorizations has demonstrated promising performance in terms of extracting and characterizing developmental brain patterns [31] as well as for the extraction of developmental features from EEG pediatric data [32] and for the study of functional connectivity patterns [33]. Many recent studies have analyzed neuroimaging data from multiple modalities using multiway representations and used CMTF-like methods to jointly analyze EEG/MEG [34], [35], EEG and fMRI [2], [17], [18]. Despite the substantial number of studies dedicated to the joint high-order analysis of EEG and fMRI, the joint tensor-based processing and fusion of MEG and fMRI have not been well investigated. We found a few studies that focused on MEG and fMRI data fusion using multivariate pattern analysis techniques such as representation similarity analysis (RSA) [36], [37], [38], the RSA-like model-based MEG and fMRI fusion [39] that used a commonality analysis [40]. The RSA method presented in [36], [37], [38] is similarity-based fusion, which utilizes a region-of-interest (ROI) approach to relate MEG and fMRI in selected ROIs by computing distance matrices between fMRI brain regions and MEG time courses. Even though similarity-based fusion presents a novel application for the fusion of MEG and fMRI, it is not a data-driven fusion method. It is instead a model-based

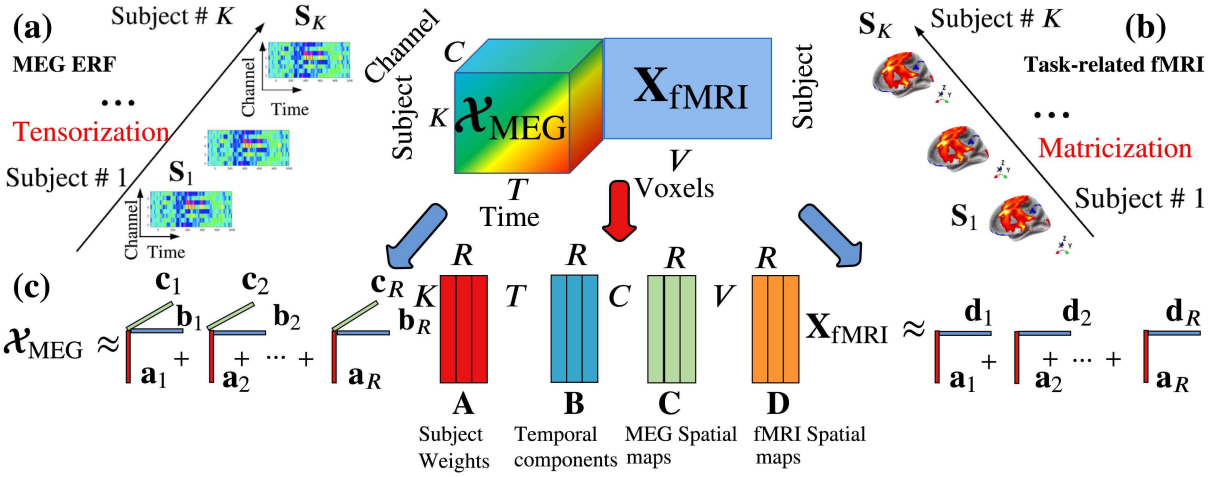

Fig. 1. Illustration of joint tensor/matrix analysis for multi-subject MEG and fMRI data. (a)–(b) Joint MEG and fMRI tensor matrix model. MEG tensor formation is achieved by arranging the subject's ERF responses along the first dimension. fMRI data are arranged as a matrix in which spatial maps derived from task-related contrast are concatenated along the first dimension. The MEG tensor is coupled with the fMRI matrix in subject mode. (c) MEG and fMRI tensor/matrix decomposition into R rank-1 components. Each rank-1 component represents a distinct spatiotemporal brain activity pattern with subject weight ($\mathbf{a}_r$), temporal ($\mathbf{b}_\mathbf{r}$), MEG spatial signatures ($\mathbf{c}_\mathbf{r}$), and fMRI spatial signatures ($\mathbf{d}_\mathbf{r}$).

method that restricts the exploration of associations between MEG and fMRI based on the prior knowledge of the activation patterns. Another limitation of the similarity-based fusion is the inherent ambiguity [3] between spatial brain regions and MEG timepoints, which results in a non-unique solution, and, thereby, limited interpretability.

To address the limitations of EEG and fMRI data fusion as well as existing model-driven methods for MEG and fMRI joint data analysis, and to facilitate unrestricted interaction between modalities while ensuring model uniqueness, we have employed a high-order decomposition known as "structure-revealing data-driven fusion" [12]. We employed the CMTF model to jointly analyze MEG event-related fields (ERF) and fMRI contrast data. MEG ERFs are organized as a tensor with dimensions of subject $\times$ time $\times$ sensors, while the fMRI data are represented as a matrix with dimensions of subject $\times$ voxels. The proposed high-order data fusion exploits multilinear relationships within MEG, enforces the uniqueness of overall decomposition, and enables unsupervised full interaction among MEG and fMRI modalities. The proposed approach is shown in Fig. 1. We quantitatively evaluated the advantages of joint tensor/matrix decomposition by comparing the discriminative abilities of the CMTF model and single modality tensor/matrix models, including the CP and multi-subject independent component analysis (ICA) [41], [42].

In this paper, we demonstrate the effectiveness of multimodal MEG and fMRI tensor/matrix decomposition for robust extraction of task-related joint latent components that can explain brain function via multiway data-driven fusion. We show that the joint latent components extracted via multimodal tensor/matrix decomposition possess higher discriminative ability to differentiate between subject subgroups than a single-modality's decomposition for tensor and matrix-based data-driven models. The proposed framework allows for the study of sensory processing that yields insights into higher-level cognitive functions and typical developmental trajectories in adolescent brains. Our results show that both fast temporal dynamics of MEG and high spatial resolution of fMRI improve the characterization of underlying neural processes by providing the spatiotemporal location of the brain activity. Therefore, joint tensor/matrix decomposition provides higher statistical power by leveraging the complementary nature of both modalities via the full interaction of MEG and fMRI data and, thus is an attractive solution for joint analysis and group-level inferences in multimodal studies.

II. Data

A. Participants

The participants included 74 healthy children and adolescents (38 males, 36 females) between the ages of 9 and 15 (mean $= 11.80$ years and std $= 1.91$), with no reported clinical diagnoses, recruited by the Mind Research Network (MRN) in Albuquerque, New Mexico as part of the Developmental Chronnecto-Genomics (Dev-CoG) study [43]. The participants and parents signed consent forms approved by the Advarra institutional review board (IRB) prior to joining the study. All procedures were approved by the IRB prior to the start of the experiment. The participants were divided into high-performance (HP, $n = 38$) and low-performance groups (LP, $n = 36$) as described in Supplementary Methods Section I-A. Participants' characteristics are shown in Supplementary Table S.I. and Fig. S.1.

B. MEG and fMRI Data Acquisition and Image Preprocessing

We provide a brief overview of the experimental and preprocessing pipeline. The detailed design and study protocol has been previously published in [31], [43]. MEG and fMRI data were separately collected from 74 subjects during a multisensory task in which the subjects were instructed to press their index finger when they saw anything, heard anything or

both. After the intertrial interval (ITI), the sensory stimulus of auditory (AUD), visual (VIS), or audio-visual (AV) was shown for 800 ms. MEG epochs around the stimulus onset were averaged across approximately 100 trials within respective conditions and formed sensor-level ERFs time-locked to the stimulus condition (AUD, VIS, or AV). For the fMRI data, we computed stimulus-related three-dimensional (3D) contrast images temporarily aligned to each target stimulus by the general-linear model using the statistical parametrical mapping software package [44].

III. Multimodal Joint Tensor/Matrix Analysis of MEG and fMRI Data

In this section, we describe the generative data model for data fusion and joint analysis of multi-subject MEG and fMRI data via joint tensor/matrix decomposition. The mathematical notations and definitions used in this paper were incorporated from [45] and [46] and are briefly described in Supplementary Methods Section I-B.

A. Generative Model for CMTF MEG and fMRI Data Fusion

According to the experimental/preprocessing paradigm described in Section II-B, we modeled the observed MEG and fMRI recordings as a linear mixture of the underlying neural sources synchronized in time across subjects within a specific target-related stimulus. We assumed that these neural sources are reflected in both modalities. Let $\mathbf{X}_{\text{MEG}}^{(k)} \in \mathbb{R}^{T\times C}$ represent the observed MEG ERF waveform for the kth subject from C common sensors synchronized across T timepoints. Additionally, $\mathbf{x}_{\text{fMRI}}^{(k)} \in \mathbb{R}^{V}$ denotes the stimulus-related fMRI spatial contrast map observed for the kth subject, where $k = 1, \ldots, K$.

We present a formal description of MEG ERF and fMRI generative data fusion model as follows. Let $\mathbf{c}_r \in \mathbb{R}^{C}$ denote the rth ERF sensor source vector, which characterizes the spatial sensitivity of each sensor to the rth source. Furthermore, the activity of the rth source at all timepoints is denoted as $\mathbf{b}_r \in \mathbb{R}^{T}$. These temporal profiles represent the level of activity exhibited by the rth source across different timepoints. Moreover, the spatial sensitivity of each fMRI spatial contrast map to the rth source is denoted by $\mathbf{d}_r \in \mathbb{R}^{V}$. Lastly, $\mathbf{a}_r$ denotes the subject weights assigned to the rth source for K subjects. The subject weights denote the contribution of the rth source to the neural activity of the kth subject.

In the absence of noise, the multisubject MEG ERFs, denoted as $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{K\times T\times C}$, and fMRI spatial contrast maps $\mathbf{X}_{\text{fMRI}} \in \mathbb{R}^{K\times V}$ observed in K participants, can be expressed as a linear instantaneous combination of the underlying sources:

$$\boldsymbol{\mathcal{X}}_{\text{MEG}} = \sum_{r=1}^{R} \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r,$$
$$\mathbf{X}_{\text{fMRI}} = \sum_{r=1}^{R} \mathbf{a}_r \circ \mathbf{d}_r. \tag{1}$$

The mixing system (1) is characterized by its associated weights. Fig. 1(a)–(b) illustrate the MEG and fMRI generative model for data fusion, where the MEG modality is encoded as a 3D tensor, and the fMRI modality is encoded as a 2D matrix, with the tensor and matrix coupled together in the common subject dimension.

B. Multi-Subject CMTF MEG and fMRI Tensor/Matrix Decomposition

By adopting the generative model shown in Fig. 1(a)–(b), the common latent structure of MEG and fMRI in (1) can be directly extracted through the CMTF [12] R-component model, which allows to exploit the multilinear structure of $\boldsymbol{\mathcal{X}}_{\text{MEG}}$.

The CMTF for joint decomposition of MEG and fMRI is formulated as the minimization of the following cost function [12]:

$$\begin{aligned} &f(\boldsymbol{\Lambda}, \boldsymbol{\Sigma}, \mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}) \\ &= \min_{\boldsymbol{\Lambda},\boldsymbol{\Sigma},\mathbf{A},\mathbf{B},\mathbf{C},\mathbf{D}} \frac{1}{2}\|\boldsymbol{\mathcal{X}}_{\text{MEG}} - [[\boldsymbol{\Lambda}, \mathbf{A}, \mathbf{B}, \mathbf{C}]]\|_F^2 \\ &+ \frac{1}{2}\|\mathbf{X}_{\text{fMRI}} - \mathbf{A}\boldsymbol{\Sigma}\mathbf{D}^{\mathsf{T}}\|_F^2, \\ &\text{s.t} \quad \|\mathbf{a}_r\|_2 = \|\mathbf{b}_r\|_2 = \|\mathbf{c}_r\|_2 = \|\mathbf{d}_r\|_2 = 1, \\ &\forall r = 1, \ldots, R, \end{aligned} \tag{2}$$

where $\|\cdot\|_F$ denotes the Frobenius norm, $\mathbf{a}_r \in \mathbb{R}^{K}, \mathbf{b}_r \in \mathbb{R}^{T}, \mathbf{c}_r \in \mathbb{R}^{C}, \mathbf{d}_r \in \mathbb{R}^{V}$ are the columns of the factor matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$ normalized to a 2-unit norm for $r = 1, \ldots, R$. The norms are absorbed into diagonal matrices $\boldsymbol{\Lambda} \in \mathbb{R}^{R\times R}$, $\boldsymbol{\Sigma} \in \mathbb{R}^{R\times R}$. The matrices $\mathbf{A} \in \mathbb{R}^{K\times R}, \mathbf{B} \in \mathbb{R}^{T\times R}, \mathbf{C} \in \mathbb{R}^{C\times R}$, $\mathbf{D} \in \mathbb{R}^{C\times R}$ correspond to the factor matrices in the subject, time, sensor, and voxel modes.

The CMTF model jointly factorizes tensor $\boldsymbol{\mathcal{X}}_{\text{MEG}}$ and matrix $\mathbf{X}_{\text{fMRI}}$ by simultaneously fitting the CP model [45] to tensor $\boldsymbol{\mathcal{X}}_{\text{MEG}}$, and factorizing matrix $\mathbf{X}_{\text{fMRI}}$ in such a way that the matrix $\mathbf{A}$ extracted from the common (subject) mode is the same for MEG and fMRI datasets. The CP model is unique under mild [15], [47] conditions. It was shown in [48] that the CMTF model inherits uniqueness from the CP model, and generates unique interpretable factors up to scaling and permutations. Therefore, using the CMTF model we can reliably relate MEG ERF and fMRI brain activity and also as examine associations between the timing of the activations, functional networks, or brain cortical regions. Of note, the component signatures $\mathbf{a}_r$, $\mathbf{b}_r$, $\mathbf{c}_r$, $\mathbf{d}_r$ are subject weights of time-varying ERF patterns, time courses, and ERF spatial and fMRI spatial contrast maps, respectively.

IV. Implementation Details

A. Data Preprocessing

Prior to the analysis, we normalized the MEG data by centering the third-order MEG tensor across the time mode and scaling within the subject mode by its standard deviation [49]. The fMRI data were centered (subject-wise), and each row was divided by its standard deviation.

B. Multimodal MEG and fMRI Component Estimation

Three MEG data tensors ($\boldsymbol{\mathcal{X}}_{\text{MEG}}^{VIS}$, $\boldsymbol{\mathcal{X}}_{\text{MEG}}^{AUD}$, $\boldsymbol{\mathcal{X}}_{\text{MEG}}^{AV}$) and three fMRI matrices were created ($\mathbf{X}_{\text{fMRI}}^{VIS}$, $\mathbf{X}_{\text{fMRI}}^{AUD}$, $\mathbf{X}_{\text{fMRI}}^{AV}$) for each of the three stimulus conditions according to the data fusion generative model. We conducted three separate CMTF decompositions for each stimulus condition with the identified tensor rank (see details on rank estimation in the Supplementary Methods Section I-C). The fitted CMTF models resulted in three estimated MEG tensors and three fMRI matrices that consisted of R-component factor matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$ that were used to reconstruct joint group-level MEG ERF and fMRI latent components. The fMRI spatial component maps were scaled to Z-values, and entered into nonparametric permutation one-tailed t-tests with a maximum t-statistics [50] and thresholded at $p < 0.05$.

C. Performance Evaluation of Multimodal MEG and fMRI Decomposition

To evaluate the performance gained by the multimodal fusion approach, we compared the sensitivity of the CMTF MEG and fMRI model for each modality with the decomposition results derived from single-modality data models. The sensitivity analyses assessed the discriminative power and robustness of the results in terms of the Cohen's d effect size (ES) [51] (see Section IV-C1) accompanied by p-values and sample size N. In particular, we compared the discriminative power gained by the MEG modality using CMTF with that of individual CP decomposition using MEG data alone. For the sensitivity comparison of the fMRI modality, we used the CMTF model and compared it with two other state-of-the-art matrix-based data-driven methods, namely, ICA using the entropy-bound rate Minimization (ERBM) algorithm [52] and the DL method [53]. The ERBM takes higher-order statistical information into account through a flexible density model as well as sample dependence for achieving ICA. On the other hand, DL provides a decomposition through a linear combination of basic elements (dictionary) through sparse coefficients by learning both matrices in an alternating fashion. For the comparison, we utilized a combined multi-subject dataset created using a generative data model based on multi-subject ICA [41], [42]. Supplementary Methods Section I-D provides additional details about the multi-subject ICA model. The components from different algorithms were matched based on the highest pairwise Pearson correlation coefficient (r) between all possible combinations of the components.

1) Statistical Group-Level Analysis: We compared the discriminative power of multimodal and single-modality models for differentiating subject subgroups using mixed analysis of covariance (ANCOVA) of component loading factors in subject mode (columns of matrix $\mathbf{A}$) [31], followed by post hoc analysis using two-tailed t-tests and corrections for multiple comparisons using the false discovery rate (FDR) [54] with the significance level of $\alpha = 0.05$ to determine statistical significance. The ANCOVAs were calculated for each component and stimulus condition while controlling for age, sex, parental socioeconomic status (SES), and subject head motion. We report F-statistics, t-values, p-values, ES, and sample size N. The ES was evaluated by the generalized η_G^2, and Cohen's d values. Additionally, we also reported the mean (M) and standard deviation (SD) of measures of interests. The thresholded ($p < 0.05$) statistical images (T-maps) and Cohen's d maps showed significant activations of fMRI spatial maps and the ES, respectively. The Cohen's d ES [55], which represent the standardized mean difference between subject subgroups was used as a performance metric for the assessment of decomposition method sensitivity as follows:

$$\text{Cohen's } d_{(r)} = \frac{M(\mathbf{a}_{r_{\text{HP}}}) - M(\mathbf{a}_{r_{\text{LP}}})}{\sqrt{\frac{(n_{\text{HP}}-1)SD(\mathbf{a}_{r_{\text{HP}}})^2+(n_{\text{LP}}-1)SD(\mathbf{a}_{r_{\text{LP}}})^2}{n_{\text{HP}}+n_{\text{LP}}-2}}}, \quad (3)$$

where $M(\cdot)$ is the mean and standard deviation $SD(\cdot)$ of the rth column of matrix $\mathbf{A}$ for a subgroup, n_{HP}, and n_{LP} are the number of subjects for the HP and LP subgroups, respectively, $r = 1, \ldots, R$.

D. Multimodal Tensor Group-Level Analysis

This section presents a multimodal tensor-group level analysis, which extends the tensor-group level analysis for a single imaging modality introduced in [31]. Similar to the CP model, the CMTF model performs simultaneous factorization and is fully multivariate [56]. An important property of the CP and CMTF decomposition is their ability to mathematically isolate the underlying latent components in the form of component (factor) matrices. Each rth latent component in the sensor space or voxel space is represented as the outer product of R components ($\mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r$) or ($\mathbf{a}_r \circ \mathbf{d}_r$), respectively, with its signatures corresponding to the subject (magnitude of the component), temporal, and spatial factors.

The resulting component matrices can be utilized in statistical tests following the joint factorization of the MEG and fMRI data using the CMTF model. Due to the fully multivariate nature of the CMTF model, each factor of the latent joint component is identified across all levels of the other factors [56]. By quantifying the magnitude of each CMTF component at every timepoint, sensor, and voxel, there is no need to selectively choose specific timepoints, sensor sites, or brain regions for extracting group magnitudes in group-level inferences. This characteristic allows for the direct application of statistical inference to the selected component signatures. To identify discriminative components that differentiate between subject subgroups (HP vs. LP), the rth column of the factor matrix $\mathbf{A}$, along with the subject subgroup, is submitted to ANCOVAs and two-tailed t-tests. The multimodal tensor group-level analysis is summarized in Fig. 2 and Table S.II, while the details of the group-level statistical inference are given in Supplementary Methods Section I-E.

1) Execution Details: We applied alternating least squares (ALS) algorithm to estimate the factor matrices for the CMTF and CP models [45]. The CMTF-ALS [12] algorithm was used to fit the CMTF model (2). For the CP model, we used the CP-ALS algorithm [45]. All models were fit 100 times with random initialization, and the most stable run was selected (see Supplementary Methods Section I-C). The CMTF and CP decomposition model order was evaluated using the core consistency diagnostic (CONCORDIA/CCD) [57], average congruence product [58], and Bayesian information criterion [59]

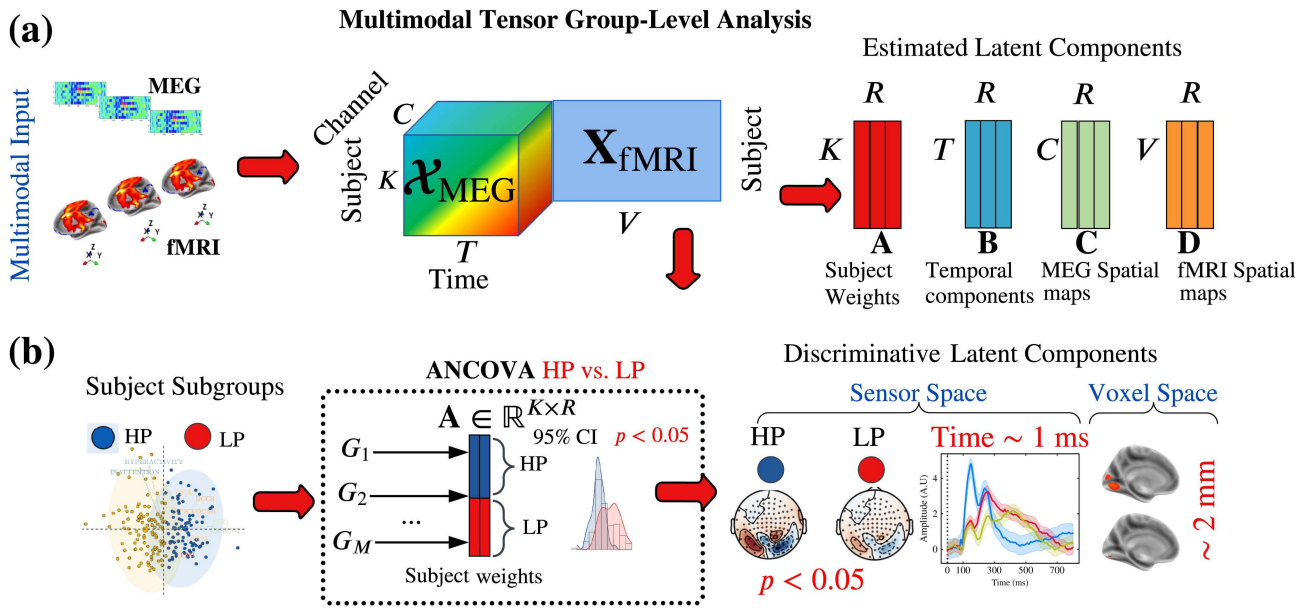

Fig. 2. Illustration of multimodal group-level tensor analysis. (a) Component matrices obtained through the CMTF decomposition represent the estimated latent components. (b) Multimodal group-level tensor analysis utilizes the component loading factors in the subject mode (columns of the factor matrix $\mathbf{A}$) as input to ANCOVA to identify discriminative latent components. Group-discriminative joint latent components are determined by statistically comparing the component loading factors of subject subgroups, aiming to identify significant between-group differences with $p < 0.05$.

TABLE I
SUMMARY OF JOINTLY ESTIMATED MEG ERF AND FMRI COMPONENTS

	MEG ERF and fMRI Joint Component			
Stimuli condition	M150/VIS	M200/mSFG	M300/AUD	M400/SM
VIS	√	√	-	√
AUD	-	√	√	√
AV	√	√	√	√

(Supplementary Methods Section I-C, Fig. S.2: VIS: $R = 3$; AUD: $R = 3$, AV: $R = 4$). The multi-subject ICA model order was determined using the minimum description length criterion [60] (VIS: $R = 3$; AUD: $R = 3$, AV: $R = 4$). Ten-fold cross-validation was used to assess the value of the regularization parameter, which was set as $\lambda = 10^{-4}$ for the DL method.

V. RESULTS

A. Multimodal Tensor Analysis Using CMTF MEG and fMRI Model

In this section, we present the results of the CMTF (2) model to jointly decompose MEG and fMRI data from multisensory tasks. The tensor analysis revealed a combined total of 10 joint components across all conditions according to selected tensor rank (VIS: $R = 3$, AUD: $R = 3$, AV: $R = 4$) (see Table I). The components were characterized on the basis of the prominent ERF peaks and fMRI activations. We present joint components extracted from the CMTF decomposition in Fig. 3, and for the VIS, AUD, and AV conditions in Figs. S.7– S.8. In Figs. 3, S.7, and S.8, every subfigure depicts fMRI component activations along with their corresponding ERF component. These ERF components display signal traces from all individual MEG sensors averaged across subject ERF components. The timecourse of the average stimulus-related ERF is depicted in yellow, while the average ERF component (averaged across sensors) is plotted in cyan. The locations of the fMRI activations in the brain are given in Montreal Neurological Institute (MNI) coordinates, and anatomical areas are described.

The VIS/M150 component was found under the VIS and AV conditions and is shown in Fig. 3(a) and Fig. S.7(a). This component was associated with the first prominent visual peak at 149–151 ms. Figs. 3(a) and S.7(a) depict the fMRI one-tailed T-map (nonparametric permutation t-test, thresholded with $p < 0.05$), which shows the activation in the left and right lingual (LING.L/R, (Broadman area) BA18) gyrus, right cuneus (CUNC.R, BA19) and bilateral calcarine (CAL.L/R, BA17) sulcus. The M200/mSFG component with an early peak at approximately 46–82 ms and 112 ms and a late peak at 171–260 ms was consistently found in the VIS, AV, and AUD conditions (see Figs. 3(b) and S.7(b), S.8(b)). The fMRI activation area associated with the M200 component was in the medial superior frontal gyrus (mSFG, BA09), precuneus (PCUN, BA7) and presupplementary motor area (pre-SMA, BA08). The right AUD/M300 component was found in the AUD and AV conditions and is shown in Figs. S.7(a) and S.8(c). The AUD/M300 component revealed three subcomponents: early ERFs at approximately 80–83 ms and 109–120 ms and a late auditory at approximately 312–378 ms. The associated fMRI activations for the AUD/M300 component were found in the bilateral Heschl's gyrus (HES.L/R, BA41/42) and left/right superior temporal gyrus (STG.L/R, BA22). The M400/SM component was extracted for all stimulus conditions (VIS, AUD, and AV), as shown in Figs. 3(c) and S.7(c) and S.8(d). The M400/SM component

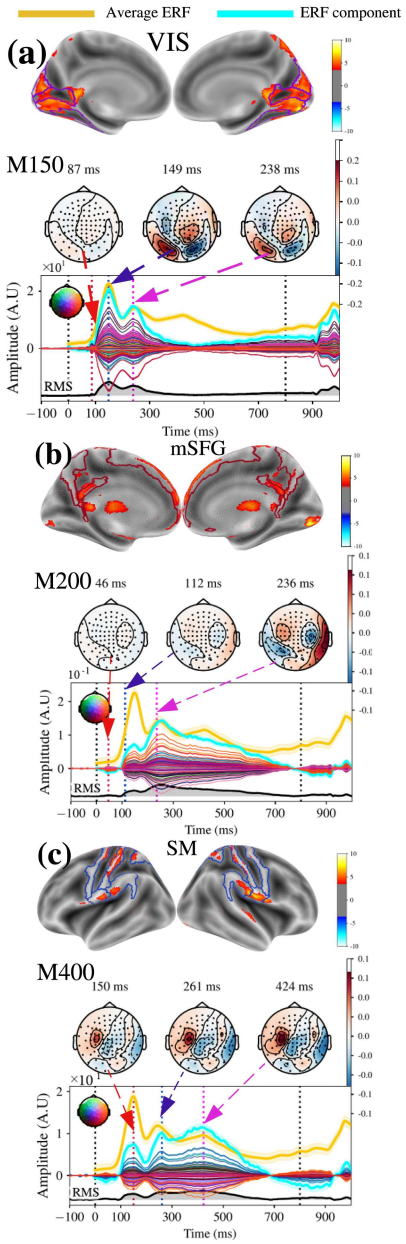

Fig. 3. Group averaged results of CMTF MEG and fMRI decomposition for the VIS stimulus. Every extracted joint component contains ERF part and an fMRI part represented as one-tailed voxelwise T-map (nonparametric permutation t-test, thresholded with $p < 0.05$). The joint components are shown in different subfigures (a)–(c). Top: fMRI component. Bottom: ERF component. The topographic maps show the density of spatial patterns that correspond to prominent time peaks denoted with red, blue and pink arrows. The average stimulus-related ERF timecourse is shown in yellow, and the average ERF (average across sensors) component is plotted in cyan. (a) The M150/VIS joint component, fMRI activation in the left and right lingual (LING) gyrus (LING/BA18 Montreal Neurological Institute (MNI) [10, −40, 0]), calcarine (CAL/BA17 MNI [−8, −78, 10]) sulcus, and cuneus (CUNC) (CUNC.R/BA19 MNI [10, −88, 34]). (b) The M200/mSFG joint component. The fMRI maps show the activity in the superior frontal gyrus (medial part) (mSFG, BA09 MNI [0, 62, 18]), presupplementary motor area (pre-SMA/BA08 MNI [16, 42, 50]), and precuneus (PCUN) (PCUN/BA07 MNI [0, −10, 6]). (c) The M400/SM joint component. The fMRI activations are in the postcentral gyrus (Post.CG/BA1/2/3 MNI [−44, 42, 60]), the precentral gyrus (primary motor cortex) (Pre.CG/BA4 MNI [−34, −28, 68]), and the bilateral superior parietal gyrus (SPG.L/R MNI [−24, −46, 72]). VIS: visual, mSFG: superior frontal gyrus (medial part), SM: sensorimotor.

consisted of early sensory subcomponents at approximately 48–150 ms and a late latency subcomponent at approximately 356–424 ms. The fMRI activations in the M400/SM component were identified in the left/right postcentral gyrus (Post.CG. L/R, BA1/2/3), left/right precentral gyrus (Pre.CG, BA4) and in the bilateral superior parietal gyrus (SPG.L/R, BA07). These fMRI activations are expected and consistent with previous findings in the literature [61]. The VIS/M150 and AUD/M300 components generated activations in the primary visual and primary auditory cortices, followed by the expected motor response (M400/SM) to press the button with the fMRI activations in the primary and somatosensory motor brain areas. The anatomical labels of the fMRI significant activation clusters are listed in Supplementary Table S.III.

B. Group-Level Discriminative Components

In this section, we studied the effect of each decomposition method (CMTF, CP, ERBM, DL) on the group-level sensitivity of estimated MEG ERF and fMRI components to differentiate between the subject subgroups.

To compare the group-level sensitivities of multimodal vs. single-modality decomposition methods, we first applied a mixed-measures four-way ANCOVA on the component loading factors for each modality, component/decomposition method, and stimulus condition after controlling for covariates. The modality (fMRI, MEG), stimulus condition (VIS, AUD, and AV) and component (VIS/M150, AUD/M300, and M200/mSFG, M400/mSFG) were within-subject factors, and the decomposition method (CMTF, CP, ERBM, and DL) and subject subgroup (HP vs. LP) were between-subject factors. The mixed-measures ANCOVA comparison of component loading factors showed a statistically significant component $\times$ decomposition method $\times$ group interaction for the VIS condition ($F_{4,648} = 14.18, p < 0.0001, \eta^2_G = 0.08$), AUD ($F_{4,643} = 13.193, p < 0.0001, \eta^2_G = 0.076$), and AV condition ($F_{6,859} = 91.596, p < 0.0001, \eta^2_G = 0.391$). After conducting post hoc analyses using independent two-tailed t-tests (FDR corrected, $p < 0.05$), we identified a total of 30 components that showed significant group differences ($N = 74$, HP vs. LP) across all decomposition methods (CMTF:10, ERBM: 10, and DL:10) and stimulus conditions (VIS, AUD, and AV). The corresponding post hoc results (HP vs. LP) are shown in Supplementary Tables S.IV and S.V. The detailed methodology of the relevant ANCOVA analyses and post hoc two-tailed t-tests is provided in Supplementary Methods Section I-E.2 and depicted in Fig. S.5.

We further quantified differences in the statistical power of decomposition methods within fMRI and MEG modality by computing the mean Cohen's d ES (3) for significant group-level components (HP vs. LP) along with the p-values (see Tables S.VI and S.VII). Fig. 4 depicts the mean ES computed for both modalities using the CMTF and state-of-the-art methods. Notably, as we can see from Fig. 4, Cohen's d ES was larger with the multimodal CMTF method than with the single-modality decomposition methods. As shown in Tables S.VI and S.VII, the p-values for the CMTF were lower compared with the CP, ERBM, and DL methods.

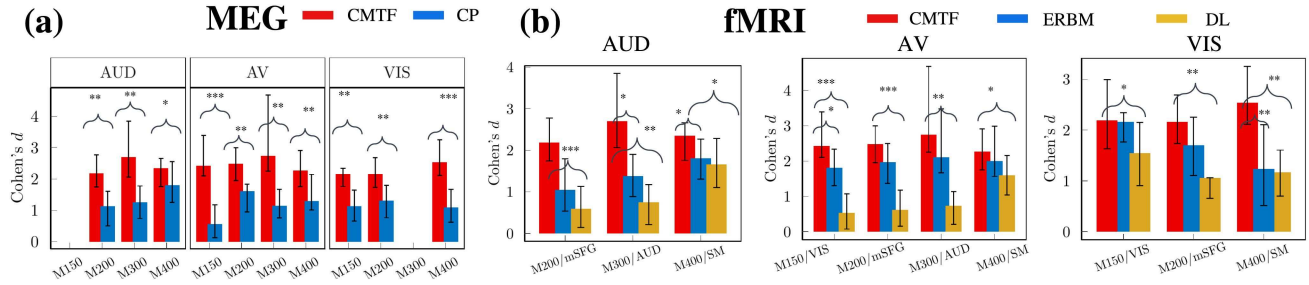

Fig. 4. Main effect of the decomposition method on the mean Cohen's d ES for the AUD, AV and VIS conditions in the MEG and fMRI modality: MEG: CMTF vs CP, fMRI: CMTF vs. ERBM vs. DL. The error bars represent the standard error of the mean. Post hoc analyses with two tailed t-tests (FDR corrected, $p < 0.05$) indicate that the mean Cohen's d ES of CMTF was significantly higher than those of the CP, ERBM and DL. The post hoc t-tests results are shown in Supplementary Tables S. IX–S.X. *: $p < 0.01$, **: $p < 0.001$, ***: $p < 0.0001$ (post hoc and FDR corrected with $p = 0.05$).

TABLE II
ONE-WAY ANCOVA (CMTF VS. ERBM VS. DL) OF COHEN'S d ES FOR THE VIS, AUD AND AV CONDITIONS IN FMRI MODALITY

Component	Cohen's d ES					
	CMTF	EBM	DL	$F(1, 146)$	p value	η^2_G
	$M(SD)$	$M(SD)$	$M(SD)$			
			VIS			
M150/VIS	2.15 (0.63)	2.18 (1.48)	1.54 (1.31)	10.85	**0.01**	0.091
M200/mSFG	2.15 (1.04)	1.70 (1.25)	1.06 (0.43)	30.01	$\mathbf{p < 0.0001}^a$	0.215
M400/SM	2.53 (1.25)	1.16 (0.98)	1.23 (1.40)	24.09	$\mathbf{p < 0.0001}^a$	0.180
			AUD			
M200/mSFG	2.18 (1.12)	1.37 (1.08)	0.74 (1.05)	41.48	$\mathbf{p < 0.0001}^a$	0.275
M300/AUD	2.71 (1.84)	1.04 (1.37)	0.59 (1.07)	54.68	$\mathbf{p < 0.0001}^b$	0.336
M400/SM	2.34 (0.98)	1.81 (1.05)	1.65 (1.28)	10.97	$\mathbf{p < 0.01}^a$	0.091
			AV			
M150	2.43 (1.38)	1.81 (1.13)	0.53 (1.08)	42.96	$\mathbf{p < 0.0001}^a$	0.282
M200	2.44 (1.14)	1.96 (1.26)	0.61 (1.12)	38.21	$\mathbf{p < 0.0001}^a$	0.259
M300	2.74 (2.28)	2.11 (1.46)	0.73 (1.01)	32.50	$\mathbf{p < 0.0001}^a$	0.231
M400	2.69 (1.26)	2.01 (1.52)	1.60 (1.22)	7.12	$\mathbf{0.002}^a$	0.061

[a] p values (FDR corrected, $p < 0.05$) from one-way ANCOVA F-tests (CMTF vs. ERBM vs. DL) adjusted for age, sex, parental SES, and subject head motion.

Consequently, to evaluate the sensitivity of the ES between methods, a one-way ANCOVA (MEG: CMTF vs. CP; fMRI: CMTF vs. ERBM vs. DL) was conducted to examine effect of the decomposition method on the mean Cohen's d ES with an FDR correction for each component (M150/VIS/, M200/mSFG, M300/AUD, M400/SM) and condition (VIS, AUD, and AV) within MEG and fMRI. There was a statistically significant difference in the mean Cohen's d ES between the CMTF and CP in MEG modality with $p < 0.05$ for nine components (Table S.VIII), and for nine components (CMTF vs. ERBM vs. DL) in fMRI with $p < 0.05$ (Table II), respectively. Post hoc analyses with pairwise two-tailed t-tests (FDR corrected, $p < 0.05$) showed several components with statistically significant method differences ($p < 0.05$) in MEG (CMTF vs. CP: Table S.IX) and fMRI (CMTF vs. ERBM vs. DL: Table S.X). The post hoc results show that the CMTF had a statistically significant higher Cohen's d ES in MEG (ES(CMTF) $>$ ES(CP)) and in fMRI (ES(CMTF) $>$ ES(ERBM); ES(CMTF) $\gg$ ES(DL); ES(ERBM) $\gg$ ES(DL)) (see Tables S.IX– S.X and Fig. 4). The detailed methodology for the sensitivity analyses using one-way ANCOVA analyses and post hoc pairwise two-tailed t-tests is provided in Supplementary Methods Section I-E.3, and depicted in Fig. S.6.

The results presented in Fig. 4 and Tables S.IX-S.X indicate that CMTF resulted in statistically significant improvements in the method sensitivity compared with the CP, ERBM, and DL methods.

1) Comparative Analysis of Decomposition Methods in fMRI Modality: We present the group-level fMRI contrast T-maps and Cohen's d maps between-subject subgroups (HP vs. LP) for the decomposition methods (CMTF, ERBM, jDL) in the VIS condition in Fig. 5. The T-maps and Cohen's d maps for the AUD and AV conditions are shown in Figs. S.9– S.10. It is evident from Figs. 4(b), S.9–S.10 that the mean values of the Cohen's d statistics were larger with the CMTF and ERBM than with the DL method. The T-maps and Cohen's d maps shown in Figs. 5 and S.9–S.10 confirm the results of the post hoc analyses with two-tailed t-tests (Table S.X: CMTF vs. ERBM vs. DL) by demonstrating higher ES for the CMTF. The second method that showed discriminative performance comparable with that of the CMTF is the ERBM, and the ES was significantly lower with the DL (see Fig. 5(b)) than with the CMTF and ERBM.

2) Comparative Analysis of Decomposition Methods in MEG Modality: We evaluated the group-level sensitivity in the MEG modality by investigating the CMTF and CP methods to differentiate the between-subject subgroups. To compare the group-level sensitivity of the CMTF and CP tensor decomposition methods, we computed timepoint/sensor-wise t-statistics from a two-tailed nonparametric permutation t-test and the mean value of Cohen's d statistics. The group-level MEG ERF components and sensor-level T-maps generated after the CMTF and CP methods in the VIS condition are shown in Fig. 6, and those for the AUD and AV conditions are shown in Supplementary Figs. S.11– S.12. We compared the sensitivities of the CMTF and CP methods using Cohen's d ES in Fig. 4(a), Figs. S.11– S.12 and Table S.IX. As shown in Fig. 4(a) and Figs. S.11– S.12, the group differences were higher and p-values lower (Tables S.VI and S.IX) with the CMTF method than with the CP method. The sensor-level T-maps shown in Figs. 6, S.11–S.12 also illustrate a higher number of significant sensors with the CMTF than with CP decomposition. Finally, the results shown in Figs. 4(a), 6, S.11–S.12 and Table S.IX show a higher statistical effect size and higher group-level sensitivity with the CMTF than with the CP decomposition method.

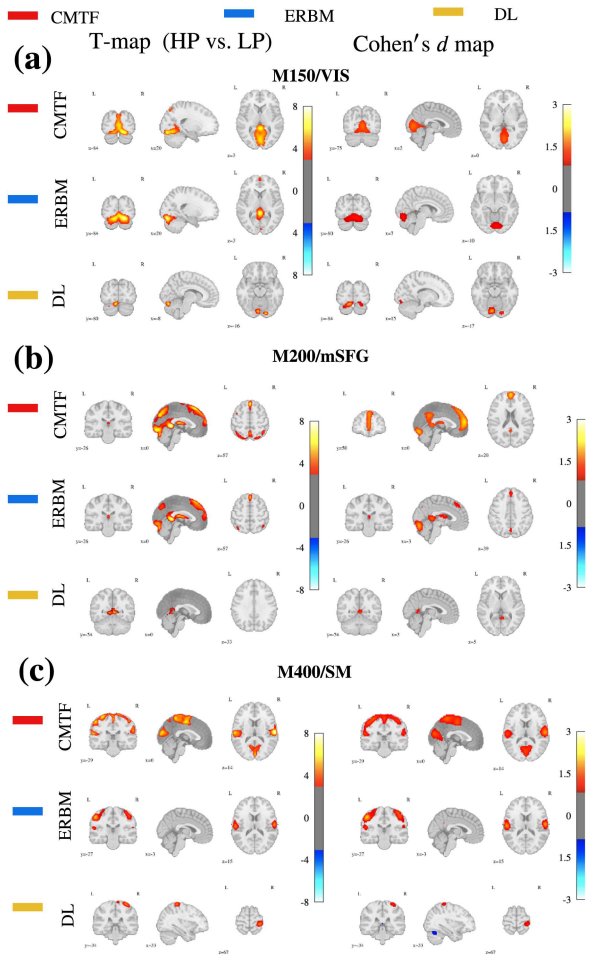

Fig. 5. Sensitivity analysis of the discriminative group-level fMRI components estimated with the multimodal CMTF vs. single-modality matrix decomposition methods for the VIS condition. The group-level contrast T-maps and Cohen's d ES maps between the HP and LP groups for the CMTF, DL, and ERBM methods are shown. Top (red): CMTF. Middle (blue): ERBM. Bottom (yellow): DL. Left: Group-level contrast T-maps (HP vs. LP). Right: Group-level Cohen's d ES maps (HP vs. LP). The T-maps and Cohen's d ES maps (nonparametric permutation two-tailed t-test with a maximum t-statistics) are thresholded at $p < 0.05$. (a) M150/VIS component. (b) M200/mSFG component. (c) M400/SM component. The post hoc two-tailed t-tests comparing the methods sensitivity using Cohen's d ES (CMTF vs. ERBM, CMTF vs. DL) are listed in Table S.X.

In summary, the results presented in Figs. 4 and S.11–S.12 show that the t-statistics and the ES were higher with the multimodal CMTF method, suggesting better sensitivity of the multimodal decomposition than of the single-modality decomposition method in both fMRI and MEG modalities.

VI. Discussion

In this paper, we jointly analysed multimodal neuroimaging data, namely, sensor-level MEG ERFs and fMRI spatial maps collected from healthy children and adolescents from our Dev-Cog study [43], using coupled tensor/matrix decomposition, such as the CMTF model. We extracted group-level ERF and fMRI latent components for the VIS, AUD, and AV stimulus conditions using the CMTF model where MEG and fMRI

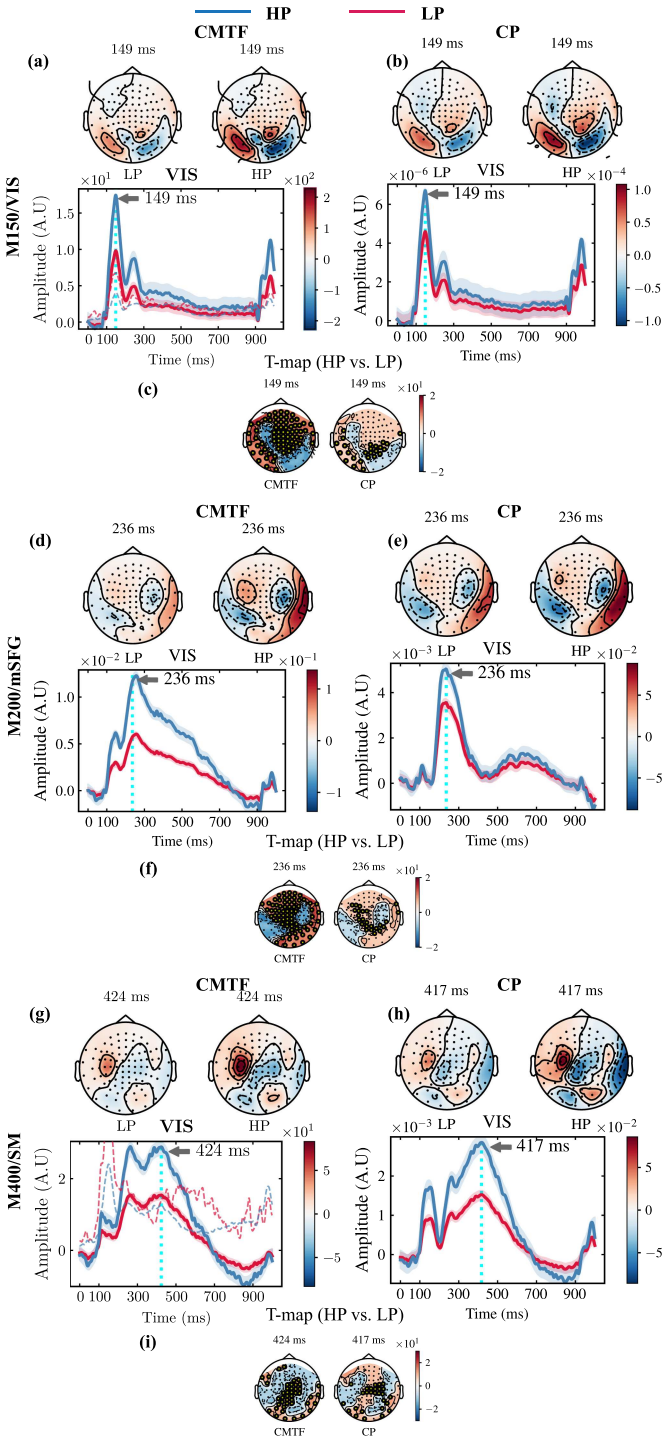

Fig. 6. Sensitivity analysis of the MEG ERF components in the VIS condition estimated different tensor decomposition methods: multimodal vs. single-modality methods. Top: Estimation of the M150/VIS, M200/mSFG, and M400/mSFG group components by the CMTF and CP decomposition methods. Bottom: The group-level T-maps between the HP and LP groups for the CMTF and CP methods are shown. The T-maps (nonparametric permutation two-tailed t-test with a maximum t-statistics thresholded at $p < 0.05$. The yellow circles on the scalp map show the locations of the significant sensors. Left: ERFs and T-maps estimated by the CMTF method. Right: ERFs estimated by the CP method. (a)–(b) M150/VIS ERF component. (d)–(e) M200/mSFG ERF component. (g)–(h) M400/SM ERF component. The post hoc two-tailed t-tests comparing the methods sensitivity using Cohen's d ES (CMTF vs. ERBM, CMTF vs. DL) are listed in Table S.IX.

were joined in the common subject dimension. In addition, we presented a framework for exploiting the ultra-high temporal resolution of MEG for the spatiotemporal reconstruction of fMRI on a millisecond scale. Furthermore, we reported the sensitivity of group-level components for multimodal and single-modality data-driven decomposition methods. We showed that CMTF enables the exploitation of multilinear structure of multisensor MEG data, and can estimate biologically meaningful patterns with uniqueness guarantees on both tensor and matrix decompositions without introducing additional constraints. We demonstrated that the joint model can (1) produce robust ERF and fMRI latent components that reveal expected task activation patterns for a given experimental paradigm and (2) provide better group-level sensitivity than that for single-modality multi-subject data-driven models.

A. Multimodal Fusion for Brain Pattern Discovery

A key objective of this paper was to demonstrate the success of multimodal fusion making use of the multiway nature of MEG data for discovering spatiotemporal brain patterns resolved in both time and space, and hence to underline the promise of CMTF for further applications. In Section V-B, we presented the results of joint ERF/fMRI component extraction for different stimulus conditions (VIS, AUD, and AV). Our findings illustrated that the CMTF model can produce latent factors that represent functionally relevant ERF/fMRI components, thus revealing meaningful spatiotemporal brain patterns. Figs. 3 and S.7–S.8 illustrate a strong correspondence between the ERF components and the average ERF waveforms. Additionally, these figures show the anticipated activations observed in fMRI components. We extracted sensory components such as the M150/VIS, M300/AUD, and sensorimotor component M400/SM, followed by the button press. These sensory components were characterized by distinct brain patterns in the visual, auditory, and primary/secondary somatosensory cortices and were consistent with the experimental paradigm. For example, the M150/VIS component depicted in Fig. 3 shows activation in occipital sensors, which aligns with the fMRI activations observed in the primary and secondary visual areas (BA17/BA18). The M150/VIS component resembles the typical visual M100/P100 wave in MEG/EEG. In neuroscience, the visual component has been used to assess visual pathway functioning [62]. The M300/AUD component consisted of expected auditory M100/M200 complex (N100/P200 in EEG) and fMRI activations in primary/secondary auditory cortices (BA41/42/22), which is typically used as a marker of auditory discrimination [62]. In addition, we identified the prefrontal M200/SFG component, which exhibited activations in the parietal-frontal regions (BA9/BA7/BA08) and was associated in the literature as a marker of executive function and attention control [63]. In summary, these findings are consistent with the existing literature and provide typical components that characterize patterns of sensory processing and higher-level cognitive functions in the healthy adolescent population. The paradigm employed was intentionally not novel in order to facilitate our ability to determine if the cross-modal results were meaningful. That is, the locations and timing of the activations are well known. It is thus particularly noteworthy that despite few cross-modal constraints in the CMTF model, the fMRI activations correspond with the expected locations of the peak activations seen in the MEG data. This interpretable and expected association between the timing and location of activity provides a strong physiological confirmation for the utility of the CMTF approach for achieving multimodal integration. Another goal of this study was to explore the relationship between specific brain patterns and cognitive abilities in individuals with varying performance levels. The study discovered distinct areas of fMRI activation that differed between children in the HP and LP groups. The results indicated significant differences between the HP and LP groups in terms of 10 discriminative components with large effect sizes (Cohen's d ES > 0.8). The statistical analyses that are shown in Tables S.IV– S.V, along with the neuropsychological tests (Table S.I.), consistently revealed the superior performance of the HP group compared with the LP group across all components and cognitive measures, including Intelligence Quotient (IQ) and Hyperactivity/Inattention. The group-level discriminative components, depicted in Figs. 5 and S.9–S.10, show contrast T-maps and Cohen's d effect size maps between the HP and LP groups. The T-maps illustrated a higher magnitude and extent of spatial activation in specific regions of the brain for the HP group, whereas the Cohen's d ES maps indicated larger statistical effects. Specifically, the HP group demonstrated higher magnitudes and extent of spatial activation in specific brain regions, while the LP group consistently exhibited reduced magnitudes and spatial extent of fMRI activations. The discriminative brain regions included bilateral primary visual cortices (BA17), medial prefrontal cortex (BA09), parietal regions (BA7), primary auditory cortex (BA41/42), and primary/secondary somatosensory brain regions (BA1/2/3). Our findings align with previous studies conducted in healthy individuals and children diagnosed with Attention Deficit Hyperactivity Disorder (ADHD) using event-related fMRI designs. These previous studies reported reduced activations in the prefrontal-parietal network [64], [65], temporo-parietal [66], sensorimotor [67], and visual brain regions [68] for patients with ADHD relative to controls. It is important to note that enrolled youth were not diagnosed with an ADHD and thus this pattern may reveal a general pattern of reduced activation in individuals with subclinical levels of inattention or lower cognitive performance. Therefore, our results, utilizing tensor-based MEG and fMRI data fusion, add an important contribution to the expanding body of literature, providing valuable insights to the sensitivity MEG and fMRI fusion to variations in cognitive function in youth.

B. Role of Multimodal Data for Group-Level Sensitivity

In Section V-B, we compared the sensitivity of the multimodal CMTF decomposition in MEG and fMRI with single-modality data-driven decomposition methods. We showed that the multimodal CMTF model can improve the sensitivity of group-level analyses versus that of single-modality decomposition data models. The results demonstrated better performance of the CMTF in terms of Cohen's d ES and lower p-values in both

MEG and fMRI modalities. Compared with single-modality decomposition models, we observed that CMTF provided 1.5–2 times higher ES (see Fig. 4) and lower p-values (see Tables S.VI–S.VII). The second best-performing method was ERBM, which showed better sensitivity than the DL method. An explanation is that the ERBM algorithm employs the flexible density-matching probability model, which automatically adjusts to the underlying data distribution.

C. Limitations and Future Work

There were some study limitations that should be considered. We have used the CMTF data model, which assumes that all modalities share the same number of factors. However, because of different nature of fused data, the number of components could vary across modalities. To address this limitation, our future work may include using the advanced CMTF model [12] that allows shared/unshared factors to quantify the contribution from each modality [12]. Future promising directions also include multitask pattern learning via the representation of the multisensory task as a CTTD data model [13]. A multitask CTTD data model could further improve the uniqueness of the joint decomposition and enable the direct evaluation of the non-additive multisensory effects.

VII. Conclusion

We demonstrated that the CMTF model can be used for learning spatiotemporal brain dynamics with high temporal resolution via data fusion of hemodynamic and electrophysiological measurements. We described a multimodal data fusion framework that allows robust identification of task-related activation patterns and improves the spatiotemporal characterization of brain processes. In this paper, we demonstrated the value of merging MEG and fMRI modalities to exploit their functional resolutions. We used the proposed approach to show that the sensitivity obtained was higher using the multimodal tensor/matrix data model in group-level analyses relative to the traditional single-modality data-driven models by leveraging complementary information across MEG and fMRI modalities. To summarize, joint tensor/matrix decomposition using the CMTF model is a promising exploratory tool for multimodal neuroimaging data. This method allowed us to study the spatiotemporal dynamics of brain function and its relation to behavioral and cognitive performance. The method enables the use of complementary information from both MEG and fMRI modalities for determining the facets of sensory processing and for characterizing the typical spatiotemporal patterns in the developing brain.

References

[1] D. Lahat, T. Adali, and C. Jutten, "Multimodal data fusion: An overview of methods, challenges, and prospects," *Proc. IEEE*, vol. 103, no. 9, pp. 1449–1477, Sep. 2015.

[2] E. Acar et al., "Unraveling diagnostic biomarkers of schizophrenia through structure-revealing fusion of multi-modal neuroimaging data," *Front. Neurosci.*, vol. 13, 2019, Art. no. 416.

[3] R. M. Cichy and A. Oliva, "AM/EEG-fMRI fusion primer: Resolving human brain responses in space and time," *Neuron*, vol. 107, no. 5, pp. 772–781, 2020.

[4] J. Sui et al., "A review of multivariate methods for multimodal fusion of brain imaging data," *J. Neurosci. Methods*, vol. 204, no. 1, pp. 68–81, 2012.

[5] V. D. Calhoun et al., "Neuronal chronometry of target detection: Fusion of hemodynamic and event-related potential data," *Neuroimage*, vol. 30, no. 2, pp. 544–553, 2006.

[6] N. M. Correa et al., "Multi-set canonical correlation analysis for the fusion of concurrent single trial ERP and functional MRI," *Neuroimage*, vol. 50, no. 4, pp. 1438–1445, 2010.

[7] R. Jin et al., "Dictionary learning-based fMRI data analysis for capturing common and individual neural activation maps," *IEEE J. Sel. Topics Signal Process.*, vol. 14, no. 6, pp. 1265–1279, Oct. 2020.

[8] T. Adali, Y. Levin-Schwartz, and V. D. Calhoun, "Multimodal data fusion using source separation: Application to medical imaging," *Proc. IEEE*, vol. 103, no. 9, pp. 1494–1506, Sep. 2015.

[9] R. F. Silva et al., "Multidataset independent subspace analysis with application to multimodal fusion," *IEEE Trans. Image Process.*, vol. 30, pp. 588–602, 2021.

[10] V. Calhoun, T. Adali, and J. Liu, "A feature-based approach to combine functional MRI, structural MRI and EEG brain imaging data," in *Proc. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2006, pp. 3672–3675.

[11] E. Martĥez-Montes et al., "Concurrent EEG/fMRI analysis by multiway partial least squares," *NeuroImage*, vol. 22, no. 3, pp. 1023–1034, 2004.

[12] E. Acar et al., "Structure-revealing data fusion," *BMC Bioinf.*, vol. 15, no. 1, pp. 1–17, 2014.

[13] C. Chatzichristos et al., "Early soft and flexible fusion of electroencephalography and functional magnetic resonance imaging via double coupled matrix tensor factorization for multisubject group analysis," *Hum. Brain Mapping*, vol. 43, no. 4, pp. 1231–1255, 2022.

[14] J. D. Carroll and J.-J. Chang, "Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition," *Psychometrika*, vol. 35, no. 3, pp. 283–319, 1970.

[15] N. D. Sidiropoulos and R. Bro, "On the uniqueness of multilinear decomposition of n-way arrays," *J. Chemometrics: A J. Chemometrics Soc.*, vol. 14, no. 3, pp. 229–239, 2000.

[16] J. Jorge, W. Van der Zwaag, and P. Figueiredo, "EEG–fMRI integration for the study of human brain function," *Neuroimage*, vol. 102, pp. 24–34, 2014.

[17] S. Ferdowsi et al., "A new informed tensor factorization approach to EEG–fMRI fusion," *J. Neurosci. Methods*, vol. 254, pp. 27–35, 2015.

[18] Y. Jonmohamadi et al., "Extraction of common task features in EEG-fMRI data using coupled tensor-tensor decomposition," *Brain Topogr.*, vol. 33, no. 5, pp. 636–650, 2020.

[19] S. P. Ahlfors et al., "Sensitivity of MEG and EEG to source orientation," *Brain Topogr.*, vol. 23, pp. 227–232, 2010.

[20] M. Illman et al., "Comparing MEG and EEG in detecting the 20-Hz rhythm modulation to tactile and proprioceptive stimulation," *NeuroImage*, vol. 215, 2020, Art. no. 116804.

[21] S. Ferdowsi et al., "Removing ballistocardiogram artifact from EEG using short-and long-term linear predictor," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 7, pp. 1900–1911, Jul. 2013.

[22] V. Abolghasemi et al., "EEG–fMRI: Dictionary learning for removal of ballistocardiogram artifact from EEG," *Biomed. Signal Process. Control*, vol. 18, pp. 186–194, 2015.

[23] R. Leahy et al., "A study of dipole localization accuracy for MEG and EEG using a human skull phantom," *Electroencephalogr. Clin. Neuriophysiol.*, vol. 107, no. 2, pp. 159–173, 1998.

[24] R. Hari and R. Salmelin, "Magnetoencephalography: From squids to neuroscience: Neuroimage 20th anniversary special edition," *Neuroimage*, vol. 61, no. 2, pp. 386–396, 2012.

[25] S. Baillet, "Magnetoencephalography for brain electrophysiology and imaging," *Nature Neurosci.*, vol. 20, no. 3, pp. 327–339, 2017.

[26] M. Hämäläinen et al., "Magnetoencephalography–theory, instrumentation, and applications to noninvasive studies of the working human brain," *Rev. Modern Phys.*, vol. 65, no. 2, 1993, Art. no. 413.

[27] M. Antonakakis et al., "The effect of stimulation type, head modeling, and combined EEG and MEG on the source reconstruction of the somatosensory P20/N20 component," *Hum. Brain Mapping*, vol. 40, no. 17, pp. 5011–5028, 2019.

[28] S. R. McWhinney et al., "Asymmetric weighting to optimize regional sensitivity in combined fMRI-MEG maps," *Brain Topogr.*, vol. 29, pp. 1–12, 2016.

[29] M. Bullock et al., "Artifact reduction in simultaneous EEG-fMRI: A systematic review of methods and contemporary usage," *Front. Neurol.*, vol. 12, 2021, Art. no. 622719.
[30] H.-J. Lee et al., "Ballistocardiogram suppression in concurrent EEG-MRI by dynamic modeling of heartbeats," *Hum. Brain Mapping*, vol. 43, no. 14, pp. 4444–4457, 2022.
[31] I. Belyaeva et al., "Multi-subject analysis for brain developmental patterns discovery via tensor decomposition of MEG data," *Neuroinformatics*, vol. 21, no. 1, pp. 115–141, 2023.
[32] E. Kinney-Lang et al., "Introducing the joint EEG-development inference (JEDI) model: A multi-way, data fusion approach for estimating paediatric developmental scores via EEG," *IEEE Trans. Neural Syst. Rehab. Eng.*, vol. 27, no. 3, pp. 348–357, Mar. 2019.
[33] Y. Zhu et al., "Discovering dynamic task-modulated functional networks with specific spectral modes using MEG," *NeuroImage*, vol. 218, 2020, Art. no. 116924.
[34] K. Naskovska et al., "Coupled CP decomposition of simultaneous MEG-EEG signals for differentiating oscillators during photic driving," *Front. Neurosci.*, vol. 14, 2020, Art. no. 261.
[35] Y. Cheng et al., "Using PARAFAC2 for multi-way component analysis of somatosensory evoked magnetic fields and somatosensory evoked electrical potentials," in *Proc. IEEE 10th Sensor Array Multichannel Signal Process. Workshop*, 2018, pp. 385–389.
[36] R. M. Cichy et al., "Similarity-based fusion of MEG and fMRI reveals spatio-temporal dynamics in human cortex during visual object recognition," *Cereb. Cortex*, vol. 26, no. 8, pp. 3563–3579, 2016.
[37] S.-M. Khaligh-Razavi et al., "Tracking the spatiotemporal neural dynamics of real-world object size and animacy in the human brain," *J. Cogn. Neurosci.*, vol. 30, no. 11, pp. 1559–1576, 2018.
[38] Y. Mohsenzadeh et al., "Reliability and generalizability of similarity-based fusion of MEG and fMRI data in human ventral and dorsal visual streams," *Vision*, vol. 3, no. 1, 2019, Art. no. 8.
[39] M. N. Hebart et al., "The representational dynamics of task and object processing in humans," *Elife*, vol. 7, 2018, Art. no. e32816.
[40] D. R. Seibold and R. D. McPHEE, "Commonality analysis: A method for decomposing explained variance in multiple regression analyses," *Hum. Commun. Res.*, vol. 5, no. 4, pp. 355–365, 1979.
[41] V. D. Calhoun et al., "A method for making group inferences from functional MRI data using independent component analysis," *Hum. Brain Mapping*, vol. 14, no. 3, pp. 140–151, 2001.
[42] V. D. Calhoun and T. Adali, "Multisubject independent component analysis of fMRI: A decade of intrinsic networks, default mode, and neurodiagnostic discovery," *IEEE Rev. Biomed. Eng.*, vol. 5, pp. 60–73, 2012.
[43] J. M. Stephen et al., "The developmental chronnecto-genomics (Dev-CoG) study: A multimodal study on the developing brain," *NeuroImage*, vol. 225, 2021, Art. no. 117438.
[44] *SPM12*, 2022. Accessed: Oct. 17, 2022. [Online]. Available: https://www.fil.ion.ucl.ac.uk/spm/software/spm12/
[45] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, 2009.
[46] A. Cichocki et al., "Tensor networks for dimensionality reduction and large-scale optimization: Part 1 low-rank tensor decompositions," *Found. Trends Mach. Learn.*, vol. 9, no. 4/5, pp. 249–429, 2016.
[47] J. B. Kruskal, "Three-way arrays: Rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics," *Linear Algebra Appl.*, vol. 18, no. 2, pp. 95–138, 1977.
[48] M. Sørensen and L. D. De Lathauwer, "Coupled canonical polyadic decompositions and (coupled) decompositions in multilinear rank-(L_r,n,l_r,n,1) terms–Part I: Uniqueness," *SIAM J. Matrix Anal. Appl.*, vol. 36, no. 2, pp. 496–522, 2015.
[49] R. Bro and A. K. Smilde, "Centering and scaling in component analysis," *J. Chemometrics*, vol. 17, no. 1, pp. 16–33, 2003.
[50] T. E. Nichols and A. P. Holmes, "Nonparametric permutation tests for functional neuroimaging: A primer with examples," *Hum. Brain Mapping*, vol. 15, no. 1, pp. 1–25, 2002.
[51] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*. Evanston, IL, USA: Routledge, 2013.
[52] G.-S. Fu et al., "Blind source separation by entropy rate minimization," *IEEE Trans. Signal Process.*, vol. 62, no. 16, pp. 4245–4255, Aug. 2014.
[53] J. Mairal et al., "Online learning for matrix factorization and sparse coding," *J. Mach. Learn. Res.*, vol. 11, no. 1, pp. 19–60, 2010.
[54] Y. Benjamini et al., "Controlling the false discovery rate in behavior genetics research," *Behav. Brain Res.*, vol. 125, no. 1/2, pp. 279–284, 2001.
[55] G. M. Sullivan and R. Feinn, "Using effect size–or why the P value is not enough," *J. Graduate Med. Educ.*, vol. 4, no. 3, pp. 279–282, 2012.
[56] F. Cong et al., "Tensor decomposition of EEG signals: A brief review," *J. Neurosci. Methods*, vol. 248, pp. 59–69, 2015.
[57] R. Bro and H. A. Kiers, "A new efficient method for determining the number of components in PARAFAC models," *J. Chemometrics: A J. Chemometrics Soc.*, vol. 17, no. 5, pp. 274–286, 2003.
[58] G. Tomasi and R. Bro, "PARAFAC and missing values," *Chemometrics Intell. Lab. Syst.*, vol. 75, no. 2, pp. 163–180, 2005.
[59] G. Schwarz, "Estimating the dimension of a model," *The Ann. Statist.*, vol. 6, pp. 461–464, 1978.
[60] Y.-O. Li et al., "Estimating the number of independent components for functional magnetic resonance imaging data," *Hum. Brain Mapping*, vol. 28, no. 11, pp. 1251–1266, 2007.
[61] B. Mijović et al., "The "why" and "how" of JointICA: Results from a visual detection task," *NeuroImage*, vol. 60, no. 2, pp. 1171–1185, 2012.
[62] C. A. Nelson and J. P. McCleery, "Use of event-related potentials in the study of typical and atypical development," *J. Amer. Acad. Child Adolesc. Psychiatry*, vol. 47, no. 11, pp. 1252–1261, 2008.
[63] R. E. Jung and R. J. Haier, "The parieto-frontal integration theory (P-FIT) of intelligence: Converging neuroimaging evidence," *Behav. Brain Sci.*, vol. 30, no. 2, pp. 135–154, 2007.
[64] M. D. Albaugh et al., "Inattention and reaction time variability are linked to ventromedial prefrontal volume in adolescents," *Biol. Psychiatry*, vol. 82, no. 9, pp. 660–668, 2017.
[65] L. Tamm et al., "Parietal attentional system aberrations during target detection in adolescents with attention deficit hyperactivity disorder: Event-related fMRI evidence," *Amer. J. Psychiatry*, vol. 163, no. 6, pp. 1033–1043, 2006.
[66] K. Rubia et al., "Temporal lobe dysfunction in medication-naive boys with attention-deficit/hyperactivity disorder during attention allocation and its relation to response variability," *Biol. Psychiatry*, vol. 62, no. 9, pp. 999–1006, 2007.
[67] S. H. Mostofsky et al., "Atypical motor and sensory cortex activation in attention-deficit/hyperactivity disorder: A functional magnetic resonance imaging study of simple sequential finger tapping," *Biol. Psychiatry*, vol. 59, no. 1, pp. 48–56, 2006.
[68] T. S. Hale et al., "Visual network asymmetry and default mode network function in ADHD: An fMRI study," *Front. Psychiatry*, vol. 5, 2014, Art. no. 81.