UCLA
UCLA Previously Published Works

Title

A study of longitudinal trends in time-frequency transformations of EEG data during a learning experiment

Permalink

https://escholarship.org/uc/item/3bw0z9x3

Authors

Boland, Joanna
Telesca, Donatello
Sugar, Catherine
<u>et al.</u>

Publication Date

2022-03-01

DOI

10.1016/j.csda.2021.107367

Peer reviewed

A study of longitudinal trends in time-frequency transformations of EEG data during a learning experiment[☆]

Joanna Boland[a], **Donatello Telesca**[a], **Catherine Sugar**[a,b,c], **Shafali Jeste**[c], **Cameron Goldbeck**[a], **Damla Senturk**[a,b,*]

[a]Department of Biostatistics, University of California Los Angeles, Los Angeles, CA 90025, USA

[b]Department of Statistics, University of California Los Angeles, Los Angeles, CA 90025, USA

[c]Department of Psychiatry and Biobehavioral Sciences, University of California Los Angeles, Los Angeles, CA 90025, USA

Abstract

EEG experiments yield high-dimensional event-related potential (ERP) data in response to repeatedly presented stimuli throughout the experiment. Changes in the high-dimensional ERP signal throughout the duration of an experiment (longitudinally) is the main quantity of interest in learning paradigms, where they represent the learning dynamics. Typical analysis, which can be performed in the time or the frequency domain, average the ERP waveform across all trials, leading to the loss of the potentially valuable longitudinal information in the data. Longitudinal time-frequency transformation of ERP (LTFT-ERP) is proposed to retain information from both the time and frequency domains, offering distinct but complementary information on the underlying cognitive processes evoked, while still retaining the longitudinal dynamics in the ERP waveforms. LTFT-ERP begins by time-frequency transformations of the ERP data, collected across subjects, electrodes, conditions and trials throughout the duration of the experiment, followed by a data driven multidimensional principal components analysis (PCA) approach for dimension reduction. Following projection of the data onto leading directions of variation in the time and frequency domains, longitudinal learning dynamics are modeled within a mixed effects modeling framework. Applications to a learning paradigm in autism depict distinct learning patterns throughout the experiment among children diagnosed with Autism Spectrum Disorder and their typically developing peers. LTFT-ERP time-frequency joint transformations are shown to bring an additional level of specificity to interpretations of the longitudinal learning patterns related to underlying cognitive processes, which is lacking in single domain analysis (in the time or the frequency domain only). Simulation studies show the efficacy of the proposed methodology.

Keywords

Event-related potentials; Longitudinal functional data analysis; Mixed effects models; Multidimensional PCA; Wavelets

1. Introduction

Autism spectrum disorder (ASD) is a heterogeneous neurodevelopmental disorder characterized by social interaction and communication impairments. Our motivating study was conducted at UCLA by our collaborator Dr. Shafali Jeste on implicit learning of children with ASD and their typically developing (TD) peers (Jeste et al., 2015). Implicit learning is defined by the detection of irregularities in one's environment without a conscious awareness or intention to learn. Children two to five years old were shown a continuous stream of colored geometrical shapes on a computer screen (see Fig. 1(a)). The shapes were presented in pairs and the children were expected to learn the order within the shape pairs as the experiment progressed. This constituted implicit learning, providing insights into core cognitive deficits and social behavior of the children in the two diagnostic groups (ASD vs. TD). As learning was expected to take place over the course of the experiment, in response to the repeated presentation of stimuli (colored geometrical shapes), referred to as trials, capturing longitudinal changes in electroencephalography (EEG) signals over trials was one of the main goals of the experiment.

EEG is a non-invasive and widely-available (low cost) brain imaging modality which records electrical activity in the brain. An event-related potential (ERP) is defined as the EEG waveform measured in response to presentation of each stimulus (e.g. a colored geometrical shape) in an EEG experiment. Analysis of EEG and ERP data dates back to 1950's in a wide spectrum of biomedical applications including epilepsy, sleep disorders, multiple sclerosis, brain tumors, lesions, major affective disorder, schizophrenia, alcoholism, bipolar mood disorder, assessment of surgical outcomes, confirmation of brain death, and clinical trials for drug development (Gasser and Molinari, 1996; Tierney et al., 2012). Typical analysis of ERP data averages the ERP signal over all the trials of the experiment, enhancing the signal-to-noise ratio (SNR) (Gasser and Molinari, 1996; Delorme and Makeig, 2004; Tierney et al., 2012). While this common technique is effective in increasing the SNR of the ERP data, it collapses information gained during the course of the experiment. This longitudinal information is important, especially in learning experiments, where it characterizes the learning trends across the study participants, including speed of learning. Previous works have been proposed to study the longitudinal changes over the course of a learning experiment. Hasenstab et al. (2015) proposed the moving average preprocessed ERP (MAP-ERP) which averages ERPs over trials in a sliding-window to retain the inherent longitudinal information. Fiecas and Ombao (2016) proposed to study the longitudinal evolution of learning via the use of time-varying spectral densities in the frequency domain. Additional frequency domain approaches include Motta and Ombao (2012) using evolutionary factor analysis to study the multi-channel EEG dynamics across trials of a motor-visual task. For frequency domain analysis of EEG data from multiple

subjects, see Krafty et al. (2011) and Krafty et al. (2017), where covariate effects on the power spectra of multiple time series are modeled in sleep studies.

The two previously proposed approaches of Hasenstab et al. (2015) and Fiecas and Ombao (2016); Ombao and Ho (2006) for studying longitudinal trends in EEG experiments analyze ERP in different domains; in the time domain and the frequency domain, respectively. While the time domain analysis of ERP data concentrates on interpretations of the commonly studied ERP phasic components, such as the P3 as shown in Fig. 1(c), the frequency domain analysis concentrates on interpretations of power from different frequency bands: delta (1–4 Hz), theta (4–8 Hz), alpha (8–16 Hz), beta (16–32 Hz) and gamma (over 32 Hz). In our motivating implicit learning paradigm, N1 and P3 are the phasic components typically observed in the ERP waveforms. The N1 dip, with a short latency (time-delay), is thought to be related to early category recognition, while the P3 peak, with a long latency, is traditionally related to cognitive processes such as signal matching, decision making and memory updating (Bugli and Lambert, 2006; Jeste et al., 2015). In the frequency domain, frequencies in the delta and theta bands have been reported to contribute to a P3 phasic component where frequencies in the delta band are associated with evaluative cognitive processing, and in the theta band are associated with the orienting response to novel stimuli (Bernat et al., 2007; Harper et al., 2014). Since both time and frequency domain analysis of ERP data carry different but complimentary information about the observed signal, we consider a time-frequency decomposition of the ERP data, targeting even richer information than is available in single domain analysis. We propose the longitudinal time-frequency transformation of ERP (LTFT-ERP) method, where a wavelet transformation is applied to the ERPs. LTFT-ERP, not only targets richer information in the signal through time-frequency transformations, it also allows modeling of longitudinal changes in the signal over trials throughout the learning experiment, adding an additional dimension for analysis (referred to as the longitudinal dimension).

Bernat et al. (2005, 2007) proposed time-frequency transformations (TFTs) of ERP data via wavelets and further included dimension reduction of the high dimensional time-frequency power surfaces through principal components analysis (PCA). To capture the longitudinal changes throughout the learning experiment, the resulting data from the proposed LTFT-ERP is even higher dimensional in our applications since the TFTs are repeated over multiple trials of the experiment. To adopt a data-driven approach to dimension reduction, similar to Bernat et al. (2005), we employ a multidimensional principal component analysis (MDPCA). We characterize vectorized TFT power surfaces as the functional dimension of the data and repetitions over trials as the longitudinal dimension. Under the simplifying assumption that the direction of variation in the functional dimension of the data stays the same for fixed slices along the longitudinal dimension, the eigenvectors in the functional dimension are obtained. Projections of the data onto the leading eigenvectors in the functional dimension allow us to study longitudinal changes in the resulting PCA scores via a mixed effects model. Finally, diagnostic-group level inference on longitudinal trends, representing learning dynamics, is derived via mixed effects modeling machinery.

The paper is organized as follows. Section 2 outlines the proposed LTFT-ERP approach, including dimension reduction via MDPCA following the multidimensional

TFT decompositions and modeling of the longitudinal trends via a mixed effects model. Simulation studies to study the efficacy of the LTFT-ERP in modeling of longitudinal trends in the MDPCA scores are outlined in Section 3, followed by applications to the implicit learning paradigm in Section 4. We conclude with a brief discussion given in Section 5.

2. The proposed longitudinal time-frequency transformation of ERP data (LTFT-ERP)

The proposed longitudinal time-frequency transformation (LTFT-ERP) starts with transformation of the ERP waveforms from each trial into the time-frequency (TFT) power surfaces using the wavelet transformation. Dimension reduction of the resulting trial specific time-frequency power surfaces is achieved by the data driven MDPCA in the third step. TFT power surfaces are first reshaped into two continuous dimensions (functional and longitudinal) before the application of MDPCA. The functional dimension represents wavelet power vectorized over ERP time and frequency within a trial and longitudinal dimension represents the repeatedly obtained TFT power surfaces over trials, providing the progression of the high-dimensional process throughout the experiment. Projection of the two-dimensional process onto the leading functional principal component vectors leads to the longitudinally estimated PCA scores which summarize changes in the signal over the course of the experiment. In a final step, longitudinal trends in the PCA scores are compared across participants in the two diagnostic groups (ASD vs. TD), leading to insights in speed and nature of learning. The main steps of LTFT-ERP are depicted in Fig. 2. We further expand on each step of the proposed LTFT-ERP algorithm in the following subsections.

Step 1: Multidimensional time-frequency transformation (MTFT-ERP) utilizing wavelets,

Step 2: Reshaping of TFT power surfaces into vectors,

Step 3: Dimension reduction via multidimensional PCA (MDPCA),

Step 4: Modeling of longitudinal trends in MDPCA scores via mixed effects modeling,

2.1. MTFT-ERP utilizing wavelets

The ERPs waveforms are transformed into TFT power surfaces in the first step using the wavelet function which is a simple oscillating amplitude waveform that is localized in time. Let $W_{ijs\ell}(u)$ denote the micro-voltage of the ERP for subject i, $i = 1, \ldots, N$, from electrode j, $j = 1, \ldots, J$, on trial s, $s \in S_i$, in condition ℓ (expected/unexpected), $\ell \in L_{is}$, observed at time u, $u = 1, \ldots, U$, where N, J, U denote the total number of subjects, electrodes, time points within a trial, respectively, and S_i and L_{is} denote the sets of non-missing trials and conditions at trial s for subject i, respectively. The maximum number of conditions per trial, denoted by L, equals two in our application (expected vs. unexpected). The minimum and maximum possible number of trials per subject are denoted by s_{min} and S, respectively. In addition, subjects may be partitioned into multiple diagnostic groups (e.g. TD vs. ASD), but we omit additional subscripts denoting diagnostic groups for simplicity of notation.

For the wavelet transformation, a general form of the wavelet function is selected and is referred to as the mother wavelet. The mother wavelet is then systematically stretched and contracted in time with a set of "daughter wavelets" (Bernat et al., 2005). The daughter wavelets are generated by the mother wavelet through scaling (frequency) and translation (time) parameters. The continuous and complex-valued Morlet wavelet,

$$\psi(u) = \pi^{-1/4} e^{i\omega u} e^{-u^2/2}, \tag{1}$$

commonly-used in the decomposition of ERP waveforms, is selected as the mother wavelet (Torrence and Compo, 1998). The "angular frequency", denoted by ω in (1), is set to 6, following previous literature (Farge, 1992). For a given scale/frequency $a > 0$ and translation/time parameter $b \in \mathbb{R}$, the resulting augmented daughter wavelet used in transformations is equal to $\psi\{(u-b)/a\}$. The wavelet transformation is then given as:

$$C_{ijs\ell}(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} W_{ijs\ell}(u) \bar{\psi}\left(\frac{u-b}{a}\right) du,$$

where $C_{ijs\ell}(a,b)$ denotes the wavelet coefficients which are continuous and complex, and $\bar{\psi}\{(u-b)/a\}$ is the complex conjugate of the daughter wavelet. The power, denoted by $\widetilde{X}_{ijs\ell}(a, b)$, is calculated as the squared magnitude of the wavelet coefficients, i.e. $\widetilde{X}_{ijs\ell}(a, b) = \left|C_{ijs\ell}(a, b)\right|^2$.

For a selected scale/frequency a_f, $f = 1, \ldots, F$, and translation/time parameters b_d, for $d = 1, \ldots, D$, the TFT power surfaces, $\widetilde{X}_{ijs\ell}$, $F \times D$ matrices with elements, $\widetilde{X}_{ijs\ell}(a_f, b_d)$, are vectorized into

$$\begin{aligned}\mathbf{X}_{ijs\ell} &\equiv \text{Vec}\left(\widetilde{X}_{ijs\ell}^{\text{T}}\right) = \left\{\widetilde{X}_{ijs\ell}(a_1, b_1), \ldots, \widetilde{X}_{ijs\ell}(a_1, b_D), \ldots, \widetilde{X}_{ijs\ell}(a_F, b_1), \ldots, \widetilde{X}_{ijs\ell}(a_F, b_D)\right\}^{\text{T}}, \\ &= \left\{X_{ijs\ell}(t_1), \ldots, X_{ijs\ell}(t_m)\right\}^{\text{T}},\end{aligned}$$

where the indices $t_1, \ldots, t_m$ are interpreted as the "functional" dimension of the resulting process, encompassing both ERP time and frequency components ($m = F \times D$). The TFT power surfaces, vectorized in $\mathbf{X}_{ijs\ell}$ are observed repeatedly over trials s. Interpreting s as a coarse time scale, we refer to it as the "longitudinal" dimension. Hence we conceptualize the TFT dynamic of the original signal as a set of random quantities varying over both a functional and a longitudinal dimension. In the next section we outline how this random object is further reduced in dimensionality through data-driven MDPCA.

Note that the tuning parameters F and D utilized in the wavelet transformations determine the resolution of the time-frequency transformation. While larger values of F and D allow for a higher resolution decomposition in frequency and time, respectively, these values are bounded by considerations of the total number of available EEG (over subjects, electrodes, conditions and trials) for the MDPCA decompositions. More specific guidance is provided between the ratio of the number of TFT power surfaces included in estimation of the trial-

specific covariances and $m = F \times D$, to ensure stability of the proposed PCA decompositions, in the next section.

2.2. Dimension reduction via MDPCA

For dimension reduction of the TFT power surfaces obtained in the previous section, we borrow ideas from marginal functional PCA (FPCA) (Park and Staicu, 2015). Marginal FPCA of longitudinally observed functional data relies on the assumption that the direction of variation, not the covariance itself, in the functional dimension of the data stays the same for fixed slices along the longitudinal dimension. This assumption, if assumed along both functional and longitudinal directions, called weak-separability (Chen and Müller, 2012; Chen et al., 2017), is weaker than the commonly assumed strong separability in high-dimensional data settings, which implies constant covariance along one dimension for fixed values of the other dimensions of the data. When applied to our two-dimensional TFT power surfaces, the assumption of constant direction of variation implies that the direction of variation along the functional dimension of ERP time and frequency (denoted by t) in the TFT power surfaces, stays the same across trials in the longitudinal dimension, denoted by s. This leads to the construction of a marginal covariance in the functional dimension, evaluated as an average of all functional covariances obtained at fixed trials, in estimation of the common functional directions of variation, captured by the functional eigenvectors.

For each trial s, we define the trial-specific functional covariance as $\Sigma_s := Cov(\mathbf{X}_{ijs\ell}) \in \mathbb{R}^{m \times m}$. Estimation of this quantity hinges on a moving window estimator, borrowing information across adjacent trials. Let A_s represent overlapping sets of trials of varying lengths for trials $s = s_{min},\dots,S$, with the maximum number of trials within a set denoted by k,

$$
A_s = \begin{cases} [s_{min}, 2s - s_{min}], & s < \frac{k}{2} \\ [s - \frac{k}{2} + 1, s + \frac{k}{2}], & \frac{k}{2} \leq s \leq S - \frac{k}{2} \\ [2s - S, S], & s > S - \frac{k}{2}. \end{cases}
$$

For a specific trial s, the estimator of Σ_s pools all vectors $\mathbf{X}_{ijs'\ell}$, s.t. $s' \in A_s$. This set of TFT power vectors is then centered by subtracting a mean vector $\overline{\mathbf{X}}_s \in \mathbb{R}^m$, obtained averaging across subjects i, electrodes j, conditions ℓ and trials in A_s. We denote the mean centered vectors as $\mathbf{X}_{ijs'\ell}^c := \mathbf{X}_{ijs'\ell} - \overline{\mathbf{X}}_s$. The ensuing estimator for the trial-specific covariance is defined as:

$$
\widehat{\Sigma}_s = \frac{1}{N_s - m} \sum_{i=1}^{N} \sum_{j=1}^{J} \sum_{s' \in A_s} \sum_{\ell \in L_{is'}} \mathbf{X}_{ijs'\ell}^c \mathbf{X}_{ijs'\ell}^{c\ \mathrm{T}}. \tag{2}
$$

Note that in the formulation above, TFTs are merged across diagnostic groups, scalp sections and conditions, in targeting trial-specific covariances. This merge requires that the direction of variation (captured by the estimated eigenvectors of the subsequent marginal

covariance) is similar across different grouping of TFTs (determined by diagnostic groups, conditions and scalp sections). Supplementary Materials Figures S.1–S.12 display the similarity across eigenvectors of marginal covariances targeted within the eight subgroups of subjects (two diagnostic groups, two conditions and two scalp sections) in our data application, justifying the merging of TFTs used in trial-specific covariances. Hence, the total number of TFT power vectors included in estimation of Σ_s is $N_s = J \times (\sum_i L_{is'}) \times |A_s|$, where $|A_s|$ denotes the number of elements in set A_s. The moving window parameters s_{min}, S and k, defining the number of trials included in each sliding window in A_s (ranging from 1 to k), are selected to maintain a 5:1 ratio between the number of TFT power surfaces included in estimation of Σ_s and $m = F \times D$. This procedure is recommended to ensure stability of the proposed PCA decompositions.

Given trial-specific covariances $\widehat{\Sigma}_s$, an estimator, $\widehat{\Sigma}$, of the functional marginal covariance, is targeted by a method of moments approach by averaging all trial-specific covariances at $s = s_{min}$, ..., S. Note that when the number of trials gets large, trial-specific covariances from a smaller set of trials (than the entire trial set) may be averaged in targeting the functional marginal covariance for computational feasibility. In this setting, the invariant directions of functional variation across trials, are targeted by the estimated m by 1 eigenvectors, φ_h, $h = 1, ..., m$, of the m by m functional marginal covariance matrix,

$$\widehat{\Sigma} = \sum_{h=1}^{m} \lambda_h \phi_h \phi_h^{\mathrm{T}},$$

with $\lambda_1 \quad \ldots \quad \lambda_m$ denoting the ordered eigenvalues (Greven et al., 2010; Di et al., 2009). The decomposition is typically truncated to include the first H, $H < m$, leading eigencomponents in applications. H is selected using the elbow in the fraction of variance explained where all components are selected before a relative flattening in the plot. Figs. 4 and 5 display the relative flattening in the fraction of variance explained plots in our data application. The mean centered TFT power vectors $\mathbf{X}_{ijs\ell} - \overline{\mathbf{X}}$, where $\overline{\mathbf{X}}$ denotes the overall mean TFT power vector (averaged over all conditions, subjects, electrodes, and trials), are then projected onto the H estimated leading functional eigenvectors, to obtain the longitudinal MDPCA scores:

$$Y_{ij\ell}^h(s) = \; <\mathbf{X}_{ijs\ell} - \overline{\mathbf{X}}, \phi_h> \, ,$$

where $< \cdot, \cdot >$ denotes the inner product. We allude to the longitudinal modeling of the MDPCA scores that is outlined in the next section by representing the estimated MDPCA scores as functions of the longitudinal argument s, representing trials.

Formulation of the functional marginal covariance from trial-specific covariances utilizes the constant direction of functional variation assumption across trials in the longitudinal dimension. This assumption can be assessed via checking the degree of similarity between eigenvectors of trial-specific covariances across trials. Plots of the estimated six leading eigenvectors from trial-specific covariances evaluated at trials 20, 40 and 60 are given in

Supplementary Materials Figures S.13–S.19 in our data application. Plots display sufficient similarity across trials, signaling no violation of the assumption of constant variation across trials.

2.3. Modeling of longitudinal trends in MDPCA scores

We model the longitudinal trends in the MDPCA scores, $Y_{ij\ell}^h(s)$, corresponding to the leading H functional eigenvectors, across trials via a linear mixed effects model. Multilevel random effects at the subject, $i = 1, \ldots, N$, and electrode region, $r = 1, \ldots, R$, levels are utilized to model the dependency of the data within subjects and electrodes within a scalp region (3 scalp regions containing 4 electrodes are depicted in Fig. 1(b) for our data application, within the frontal and posterior scalp sections) where spatial correlation between electrodes may exist. Additionally, we employ spline basis functions in modeling both the fixed and random effects to represent changes in the MDPCA scores longitudinally across trials.

Let $Y_{ij(r)\ell}^h(s)$ denote the hth leading MDPCA score for subject i, electrode j within region r, and condition ℓ, targeted as a function of the longitudinal index, trials (s). We model the MDPCA scores using fixed effects and a two-level random effects structure for subject and region. In our data application, the fixed effects parameters include an intercept, trial (represented by a natural cubic B-spline with four knots), group (ASD vs. TD), scalp section (frontal vs. posterior), condition (expected vs. unexpected), as well as higher order interactions (two- to four-way) between the main effects. This leads to a total of 40 fixed effects components. Let β^h denote the 40×1 column vector of fixed effects parameters, and let 5×1 vectors b_i^h and b_{ir}^h represent subject and region-level random effects, respectively. Further, let $q_{ijs\ell}$ denote the 1×40 row vector of the fixed effects matrix Q_i, corresponding to trial s, electrode j, and condition ℓ, and $q_{ij(r)s}$ be the 1×5 row vector of the random effects matrix $Q_{ij(r)}$ corresponding to trial s and electrode j in region r. We model $Y_{ij(r)\ell}^h(s)$ by

$$Y_{ij(r)\ell}^h(s) = q_{ijs\ell}\beta^h + q_{ij(r)s}b_i^h + q_{ij(r)s}b_{ir}^h + \epsilon_{ij(r)s\ell}^h,$$

$$b_i^h \sim MVN\left(0, D_{5 \times 5}^{1h}\right),$$

$$b_{ir}^h \sim MVN\left(0, D_{5 \times 5}^{2h}\right),$$

$$\epsilon_{ij(r)s\ell}^h \sim N\left(0, \sigma_h^2\right),$$

where D^{1h} and D^{2h} represent the random effects covariance matrices at the subject and region levels, respectively (Greven et al., 2010; Di et al., 2009). In addition, $\epsilon_{ij(r)s\ell}^h$

represents the error term for the hth principal component with variance σ_h^2. This multi-level mixed effects model leads to the following covariance structure assuming the random level effects are independent of the error term

$$\mathrm{Var}\{Y^h_{ij(r)\ell(s)}\} = q_{ij(r)s}D^{1h}q'_{ij(r)s} + q_{ij(r)s}D^{2h}q'_{ij(r)s} + \sigma_h^2,$$

$$\mathrm{cov}\{Y^h_{ij(r)\ell(s)}, Y^h_{ij(r)\ell(s')}\} = q_{ij(r)s}D^{1h}q'_{ij(r)s'} + q_{ij(r)s}D^{2h}q'_{ij(r)s'}, \ \forall i, j(r), s \neq s',$$

$$\mathrm{cov}\{Y^h_{ij(r)\ell(s)}, Y^h_{ij'(r)\ell(s)}\} = q_{ij(r)s}D^{1h}q'_{ij'(r)s} + q_{ij(r)s}D^{2h}q'_{ij'(r)s}, \ \forall i, j(r) \neq j'(r), s,$$

for within region correlation and

$$\mathrm{cov}\{Y^h_{ij(r)\ell(s)}, Y^h_{ij'(r')\ell(s)}\} = q_{ij(r)s}D^{1h}q'_{ij'(r')k}, \ \forall i, j(r) \neq j'(r'), s,$$

for within subject across region correlation.

We use the same design matrix for both the subject and region random effects to reduce the complexity of the model, but these two matrices can be taken to be different. In addition, the model framework can be extended to accommodate additional covariates that can model the dependency structure of the data in further detail. The number of equispaced knots for the cubic B-splines is chosen using AIC criteria in order to obtain a sufficient degree of smoothness in the modeling of the MDPCA scores (Shi et al., 1996; Rice and Wu, 2001). The model parameters are estimated using restricted maximum likelihood (REML). See Supplementary Materials Figures S.19 for histograms of the modeled eigenscores, displaying relatively symmetric distributions in our data applications, signaling no violation of the assumed normality assumption.

3. Simulation studies

3.1. Simulation setup

We conduct simulations to study the efficacy of the LTFT-ERP algorithm in modeling the longitudinal trends of the MDPCA scores. For data generation, we utilize the mixed effects fits in modeling the leading six MDPCA scores from the delta frequency band in our data application, since the leading six eigencomponents explained a higher proportion (approximately 90%) of the variation in the data from the delta frequency band than the theta band. In assessing the efficacy of the LTFT-ERP algorithm, we utilize mean error (ME) in estimation of the TFT mean power vector and functional eigenvector, in addition to group, condition and scalp section-specific mean longitudinal trajectories of the MDPCA scores and utilize prediction error (PE) in prediction of subject and region-specific longitudinal trajectories of MDPCA scores. We consider multiple simulation scenarios at varying signal-

to-noise ratios (SNR = 0.4, 0.8 and 1.6) and sample sizes (N = 80, 160 and 320), where results are reported based on 200 Monte Carlo runs.

The data fits in modeling the six leading MDPCA score trajectories and the estimated functional eigenfunctions from the delta frequency band are utilized to simulate the underlying TFT power vectors, denoted by $\mathbf{X}_{ijs\ell}^{signal}$. Additional error is added directly to the simulated TFT vectors, to avoid performing back wavelet transformations to target raw ERPs, which requires additional phase information. In order to simulate the six leading longitudinal MDPCA score trajectories, $Y_{ij(r)\ell}^{h}(s)$, h=1, …, 6, we first simulate subject and region-specific random effects, $b_i^h \sim MVN\left(0, D^{1h}\right)$ and $b_{ir}^h \sim MVN\left(0, D^{2h}\right)$, where the variance components, D^{1h} and D^{2h}, estimated in our data from the delta frequency band, are multiplied by 0.15 to reduce variability in the score trajectories to guarantee positivity of the generated TFT power vectors. Once the random effects are simulated, we generate MDPCA score trajectories via $Y_{ij(r)\ell}^h(s) = q_{ijs\ell}\beta^h + q_{ij(r)s}b_i^h + q_{ij(r)s}b_{ir}^h$, for $i=1, …$, N subjects, $j = 1, …$, 24 electrodes, $s = 5, …$, 80 trials and $\ell \in \{1, 2\}$ (expected/unexpected) conditions, mimicking our data analysis. The design matrices for the fixed and random effects are utilized as described in Section 2.3, with half of the sample designated as belonging to the ASD sample and the other half to the TD sample and the fixed effects parameter vectors, β^h, are taken as estimated from the data fits from the delta frequency band.

The simulated true TFT power vectors are obtained via $\mathbf{X}_{ijs\ell}^{signal} = \overline{\mathbf{X}} + \sum_{h=1}^{6} Y_{ij(r)\ell}^h(s)\phi_h$, utilizing the six leading eigencomponents, where φ_h denotes the estimated functional eigenvectors and $\overline{\mathbf{X}}$ denotes the estimated TFT mean power vector in the delta frequency band. Additional error is added directly to the power vectors, to obtain the simulated TFT power vectors, $\mathbf{X}_{ijs\ell} = \mathbf{X}_{ijs\ell}^{signal} + \mathbf{X}_{ijs\ell}^{noise}$. The error, $\mathbf{X}_{ijs\ell}^{noise}$, is obtained as the TFT power vector (via a wavelet transformation, corresponding to the squared radius in the polar coordinates) of the Euclidean vector (X, Y), where coordinates X and Y are both generated from a $N(0, c^2)$ distribution. Since r^2/c will follow a χ_2^2 distribution, c is chosen to correspond to varying SNR ratios, SNR = 0.4, 0.8, and 1.6, where SNR is defined as the ratio of the standard deviation of $\mathbf{X}_{ijs\ell}^{signal}$ to the standard deviation of $\mathbf{X}_{ijs\ell}^{noise}$ (which equals $2c$). Finally, missingness is induced by randomly removing a fraction of the generated TFT power vectors by sampling with replacement from the missingess profiles of the subjects in our data set from the implicit learning paradigm.

LTFT-ERP algorithm is applied to the simulated TFT power vectors to first target the estimated TFT mean power vector, $\widehat{\overline{\mathbf{X}}}$ (averaged over all conditions, subjects, electrodes, and trials), and the leading functional eigenvector, $\hat{\phi}_1$. Projections onto the leading eigenvector, yield the estimated leading MDPCA scores, $\hat{Y}_{ij(r)\ell}^1(s)$, where the group, condition and scalp section-specific true and estimated mean trajectories are denoted by $E[Y_{ij(r)\ell}^1(s)] = q_{ijs\ell}\beta^1$ and $\widehat{E[Y_{ij(r)\ell}^1(s)]} = q_{ijs\ell}\hat{\beta}^1$, respectively, where β^1 is estimated from the implicit learning paradigm data fits in the delta frequency

and used in data generation and $\hat{\beta}^1$ is estimated from the mixed effects modeling of the simulated MDPCA scores, $\hat{Y}^1_{ij(r)\ell}(s)$. To mimic our data application mean trajectories over trials are considered in eight subgroups ($g = 1, \ldots, 8$) determined by two diagnostics group (TD vs. ASD), two conditions (expected vs. unexpected) and two scalp sections (frontal vs. posterior). The true and estimated subject and region-specific predictions are denoted by $E\left[Y^1_{ij(r)\ell}(s) \mid b^1_i, b^1_{ir}\right] = q_{ijs\ell}\beta^1 + q_{ij(r)s}b^1_i + q_{ij(r)s}b^1_{ir}$ and $\widehat{E[Y^1_{ij(r)\ell}(s) \mid b^1_i, b^1_{ir}]} = q_{ijs\ell}\hat{\beta}^1 + q_{ij(r)s}\hat{b}^1_i + q_{ij(r)s}\hat{b}^1_{ir}$, where b^1_i and b^1_{ir} are the true random effects values simulated from the above specified normal distributions in data generation and $\hat{\beta}^1$, $\hat{b}^1_i$, and $\hat{b}^1_{ir}$ are estimated based on simulated data. Note that run times for a single simulation are approximately 20, 10, and 5 minutes at sample sizes $N = 320$, 160, and 80, respectively.

The mean errors (ME) in estimation of the TFT mean power vector and the leading functional eigenvector are defined by

$$\text{ME} = \frac{\sum_{p=1}^{m}\left|\overline{\mathbf{X}}(t_p) - \widehat{\overline{\mathbf{X}}}(t_p)\right|}{\sum_{p=1}^{m}\left|\overline{\mathbf{X}}(t_p)\right|}, \text{ and ME} = \frac{\sum_{p=1}^{m}\left|\phi_1(t_p) - \hat{\phi}_1(t_p)\right|}{\sum_{p=1}^{m}\left|\phi_1(t_p)\right|},$$

respectively, for $m = F \times D = 1764$ (with $D = 63$ and $F = 28$), similar to our data application. The ME in estimation of MDPCA mean trajectories for subjects (in two diagnostic groups: TD vs. ASD), electrodes (in two scalp regions frontal vs. posterior) and conditions (expected vs. unexpected) from a total of eight unique subgroup trajectories ($\{i, j, \ell\} \in g = 1, \ldots, 8$), summed over trials $s = 5, \ldots 80$, similar to our data application, is defined as

$$\text{ME} = \frac{\sum_{g=1}^{8} \sum_{s=5}^{80} \left| E[Y^1_{ij(r)\ell}(s)] - \widehat{E[Y^1_{ij(r)\ell}(s)]}\right|}{\sum_{g=1}^{8} \sum_{s=5}^{80}\left|E[Y^1_{ij(r)\ell}(s)]\right|}.$$

The PE in prediction of the subject and region-specific MDPCA score trajectories is defined as

$$\text{PE} = \frac{\sum_{i,j,s,\ell} \left| E[Y^1_{ij(r)\ell}(s) \mid b_i, b_{ir}] - \widehat{E[Y^1_{ij(r)\ell}(s) \mid b_i, b_{ir}]}\right|}{\sum_{i,j,s,\ell}\left|E[Y^1_{ij(r)\ell}(s) \mid b_i, b_{ir}]\right|},$$

where $i = 1, \ldots, N$ denotes subjects, $j = 1, \ldots, 24$ denotes electrodes from six scalp regions, $s = 5, \ldots, 80$ denotes trials, and $\ell \in \{1, 2\}$ denotes the two conditions (expected/unexpected).

3.2. Simulation results

The medians, 10th and 90th percentiles of the ME in estimation of the TFT mean power vector and functional eigenvector, in addition to group, condition and scalp section-specific mean longitudinal trajectories of the MDPCA scores and of the PE in prediction of subject and region-specific longitudinal trajectories of MDPCA scores are given in Table 1 at

three SNRs (SNR = 0.4, 0.8, and 1.6) and three sample sizes (N = 80, 160, and 320). Table 2 reports the ME and PE of the MDPCA score trajectories separately within the eight subgroups, determined by the two diagnostic groups (TD vs. ASD), two conditions (expected vs. unexpected) and the two scalp sections (frontal vs. posterior). This provides a better understanding of the subgroup variability in these measures. In addition, Fig. 3 displays the estimated fixed effects means of the MDPCA scores from the run with the median ME value for one of the eight subgroups, ASD expected posterior, at varying SNRs and sample sizes.

Table 1 shows a clear and expected trend in the ME and PE with varying SNR values, where both metrics decrease with increasing SNR values. Note that this trend is present at all three sample sizes. For estimation of the MDPCA trajectories, ME values for the mean longitudinal trajectories are smaller than the PE values for subject and region-specific predictions, as expected, since it is harder to predict record-specific trajectories than recover group means in a mixed effects model. In addition, there is a subtle trend of decreasing ME with increasing sample size in estimation of the mean longitudinal trajectories, as expected. The ME values reported are largely stable across the three sample sizes for mean power and eigenvector estimation, perhaps due to the large total number of ERPs observed even for the lower sample size at N = 80, due to repetitions over electrodes and conditions. (The lower sample size at N = 80 is selected not only to mimic the sample size in our data application, but also to guarantee stable MDPCA decompositions.)

Table 2 highlights the variability in the above outlined trends across the eight subgroups. More specifically while the decreasing trend with increasing SNR is pronounced for the PE values consistently across all eight subgroups, they are not consistently observed across all the subgroups in the ME values (the trends with increasing SNR are stronger for ME in some subgroups, such as both the TD and ASD groups in the unexpected condition from the frontal scalp section). This is again due to the comparatively higher PE values for the harder task of prediction compared to estimation of the group means, reflected through the ME metric. In addition, there are subtle differences in the trends observed across scalp sections in the two diagnostic groups. While ME values are larger in the ASD group in the frontal section (compared to the posterior section), they are larger in the TD group in the posterior section. Note that these differences are due to the different mixed effects coefficients used in data generation (based on fits to the original data). For example, the targeted contrasts are closer to zero in the ASD group in the frontal section (compared to posterior) and the TD group in the posterior section (compared to frontal), leading to larger ME values due to the standardization by a smaller integrand in the denominator. Fig. 3 displays that the group mean estimates get closer to the true mean with increasing SNR and sample size for the MDPCA scores from the ASD group in the expected condition in the posterior scalp section. Results confirm that even at lower SNR values, where higher ME values are observed in the estimation of the mean power and eigenvector, group mean trajectories of the MDPCA scores are effectively recovered through the efficient mixed effects modeling proposed.

4. Data analysis

4.1. The implicit learning paradigm

The implicit learning paradigm included 27 children with Autism Spectrum Disorder (ASD) and 34 typically developing (TD) peers ranging in age from 2–5 years old. Children were presented with a continuous stream of six-colored shapes (pink squares, blue crosses, yellow circles, turquoise diamonds, gray triangles, and red octagons) which were grouped into three pairs. Presentation of each shape was a trial, resulting in an ERP waveform. The order within the paired shapes stayed the same throughout the experiment and hence could be learned as the experiment progressed, while the order across shape pairs was random. Therefore, the transition within a shape pair corresponded to the "expected" condition of the experiment and across shape pairs corresponded to the "unexpected" condition (L = 2, see Fig. 1(a)). Implicit learning was detected through the difference in the ERP waveforms between the expected and unexpected conditions, and the rate of change in these differences reflected the evolution of implicit learning in the different diagnostic groups (ASD vs. TD) throughout the experiment.

There were 120 shapes presented in each condition, corresponding to the 120 trials of the experiment. EEG was recorded at J = 24 electrodes located in three regions (left, central, right) within two scalp sections (frontal and posterior) (see Fig. 1(b)). In addition, EEG was sampled at 250 Hz, producing a total of U = 250 ERP time points per waveform spanning 1000 ms. For further details on the preprocessing steps of the data, the readers are deferred to Hasenstab et al. (2015). In addition to the preprocessing steps, ERP data with no variability in ERP time have been removed from 3 ASD and 2 TD children (0.13% of the total ERP records have been removed before analysis).

4.2. Data analysis and interpretations

The implicit learning paradigm lead to two well-known ERP components: N1, related to early category recognition, and P3, related to cognitive processes such as signal matching, decision making and memory updating (Bugli and Lambert, 2006; Jeste et al., 2015). In our analysis, we concentrate on modeling power in the delta (.5 to 4 Hz) and theta (4 to 8 Hz) frequency bands, specifically in the way they contribute to the P3 phasic peak. Contributions of power from these two frequency bands to P3 in learning paradigms have been associated with cognitive processing (delta), and to orienting to a novel stimulus (theta) in previous works (Bernat et al., 2005, 2007; Harper et al., 2014). A two-step filtering process was applied to the individual pre-processed ERP waveforms that were down-sampled at every fourth-time point (leading to a total of $D = 63$ translation parameters in the TFT decompositions). First, a 3rd order high-pass Butterworth filter at 1.25 Hz was applied to separate the true signal from the direct current (DC) shift observed in the data. The DC shift is a well-known phenomenon in which the electrode system measures additional voltage around 1 Hz alongside the EEG signal due to chemical reactions between the metallic surface of the electrode and the conductive gel applied to the scalp. Then, we applied high-pass and low-pass 3rd order Butterworth filters at 4 Hz to the previously filtered signal to separate the delta (.5 to 4 Hz) and theta (4 to 8 Hz) frequency bands.

Hasenstab et al. (2015) which only modeled P3 amplitudes during the paradigm showed that implicit learning (signaled by the differences in the P3 peak amplitudes between the expected and unexpected conditions) took place between trials 5 and 80, with maximal condition differentiation occurring around trial $s = 30$. Hence, we considered modeling longitudinal trends within this trial range, $s_{min} = 5$ to $S = 80$, with $k = 30$ maximal trials used in sliding windows A_s in estimation of trial-specific functional covariances. For computational efficiency, trial-specific functional covariances from every third trial were averaged in targeting the marginal functional covariance. These choices guaranteed enough total number of TFT power surfaces for stable PCA decomposition of the marginal functional covariance, where TFT decompositions involved a total of $m = D \times F = 63 \times 28$ $= 1764$ total parameters (total scale parameters in frequency were $F = 28$). A sensitivity analysis to the choice of the total number of trials in the sliding windows A_s was carried out on the final mixed effects modeling results of the derived MDPCA scores, where k values within the range [20, 40] lead to similar results.

Before obtaining the leading eigenvectors in the functional domain, we assessed whether data from different diagnostic groups, conditions, scalp regions or frequency bands can be merged in estimation of the functional marginal covariances, as well as assessing the assumption of constant direction of functional variance across trials, leading to estimation of the marginal functional covariance. While the eigenvectors showed similar directions of variation within frequency bands across the diagnostic groups, conditions and scalp sections, they were sufficiently different between the two frequency bands considered, as expected due to initial filtering of ERPs (Supplementary Materials Figures S.1–S.12). This observation led us to consider two separate functional marginal covariances and eventually two separate longitudinal analysis of the MDPCA scores in the two frequency bands, which enabled the study of group differences across the two experimental conditions (expected vs. unexpected) within each frequency band. In addition, similarity of the directions of variation, captured by the eigenvectors of trial-specific functional covariances Σ_s for trials $s \in \{20, 40, 60\}$, signaled no violation of the constant direction of functional variation assumption for the data, leading to estimation of the functional marginal covariance within each frequency band (Supplementary Materials Figures S.13–S.18).

The estimated six leading eigenvectors $\varphi_h, h = 1, \ldots, 6$, of the functional marginal covariances explained approximately 60% and 90% of the total functional variation in the theta and delta frequency bands, respectively, where all six MDPCA scores associated with the six leading functional eigenvectors were modeled through the proposed mixed effects modeling framework. The mixed effects modeling for both the delta and theta frequency bands used four knots for the spline fits for the trial effect, selected by AIC. Multilevel random effects were considered at the subject and region levels, where the lme function in the R package nlme was used to fit the models.

The estimated six leading eigenvectors of the functional marginal covariance for the delta and theta frequency bands are displayed in Figs. 4 and 5, respectively, where bounds for the P3 component, expected in the time window of [190, 350] ms, are depicted with vertical dashed horizontal lines. While the eigenvectors for the delta frequency band (Fig. 4) span longer time intervals and do not isolate to the P3 expected time window that would

allow interpretations for power contributions directly to the P3, the estimated first and third leading eigenvectors for the theta frequency band (Fig. 5) highlight power contributions directly to the P3 component from lower theta (4–6 Hz) and higher theta (5–8 Hz) frequency intervals, respectively. Figs. 4 and 5 also display contrasts for ((ASD expected - ASD unexpected) - (TD expected - TD unexpected)), based on the mixed effects modeling of the MDPCA scores, obtained by projecting data onto the estimated functional eigenvectors. More specifically, the estimated contrasts along with their the associated 95% bootstrap pointwise intervals (CIs), based on resampling from subjects with replacement, are given in solid black, while the 95% pointwise CIs based directly on the mixed effects modeling are shaded in gray. Note that while the mixed effects modeling based CIs only incorporate model based uncertainty, pointwise bootstrap CIs incorporate uncertainty from the entire LTFT-ERP algorithm including the MTFT and MDPCA decompositions, hence are wider as a result.

Also provided in Figs. 4 and 5 are the 95% simultaneous bootstrap intervals (dashed black). Simultaneous CIs are based on a double bootstrap procedure, where the standard deviation of the estimated contrast is targeted by the first layer of bootstrap and the unknown constant $c_{0.95}$ multiplying the standard deviation is obtained from the second layer. The unknown constant $c_{0.95}$ is chosen as the 95th percentile of the distribution of the supremum of the normalized deviation of the bootstrap contrast estimate from the original contrast estimate in the data (normalized by the standard deviation of the bootstrap contrast targeted by the second layer of bootstrap). (See Dong et al., 2016 for additional details.) Simultaneous confidence bands can help test for group differences in the longitudinal trends in the MDPCA scores (in the entire contrast as a function of trials) and provide insight into trials that contribute to the detected group differences. We base the interpretations below on significance of group differences in condition differentiation based on the simultaneous bootstrap CIs, which are the most conservative among the three sets of CIs provided.

Building on the interpretations from the first leading eigenvector for the theta band that relatively isolate contributions to P3 from lower theta frequencies, inference on longitudinal trends of MDPCA scores from simultaneous bootstrap CIs identify significant group differences for the ((ASD expected - ASD unexpected) - (TD expected - TD unexpected)) contrast in the posterior scalp section (Fig. 5(b)). More specifically, for the leading MDPCA scores, signaling power from lower theta frequencies, the group condition differentiation contrast is significant in the posterior scalp section, where the simultaneous bootstrap CIs do not contain zero at trials 48 through 54. Note that positive scores on PC1 signal higher theta power contributions to P3, when multiplied by a positive leading eigenvector. Fig. 6 (b) and (d) further display the fitted leading score trajectories in the two conditions for the two diagnostic groups, respectively. The condition differentiation is larger in the ASD group (higher theta power contributions to P3 in the unexpected condition) in the posterior section trials 48–54 in PC1 (where simultaneous bootstrap CI for the group condition differentiation contrast does not contain zero). Hence the significant group condition differentiation is tied to the larger condition difference in this range in the ASD group. Condition differentiation in theta power contributions to P3 in the posterior scalp sections are typically associated with sensory and visual processing where the significant group condition contrast can be

interpreted as the ASD group displays signs of visual processing in the later parts of the experiment.

Note that in the frontal scalp section, the TD group displays larger condition differentiation with higher theta power contribution to P3 in the unexpected condition (compared to the expected condition) for later parts of the experiment, after trial 60 (Fig. 6(a)), where condition differentiation typically signals orienting to the novel stimulus when observed in the frontal scalp section. However while the pointwise CIs do not contain zero at trials 63 and 64 for the group condition differentiation in the frontal scalp section, the simultaneous bootstrap CIs do contain zero and do not signal significant group differences in the front scalp section (Fig. 5(a)).

4.3. Comparison of results to analysis from the time and frequency domains

The LTFT-ERP analysis, not only allows us to study dynamic trends related to learning throughout the experiment, it also allows for added interpretation from the associated time-frequency transformations, where power contributions to specific ERP components from different frequency bands carry different interpretations related to underlying cognitive processes. To highlight the specific added interpretations gained from the LTFT-ERP with the time-frequency joint transformations, we review in this section results from the time domain only and frequency domain only analysis.

Note that the prior longitudinal modeling of the implicit learning paradigm by Hasenstab et al. (2015) is carried out strictly in the ERP time domain, where differences in the longitudinal trends of the amplitude of the P3 component across trials in the frontal scalp region are depicted via a mixed effects model. Hasenstab et al. (2015) detected condition differentiation in both diagnostic groups around trial 30 of the experiment with different directions of differentiation. More specifically, while the ASD group exhibited higher P3 amplitude in the expected condition, the TD group exhibited higher P3 amplitude in the unexpected condition (results based on pointwise bootstrap CIs). In a parallel analysis of longitudinal trends in the frequency domain, we found that relative delta and theta power stay relatively constant across trials and do not display differences across conditions for both diagnostic groups.

While the time domain only analysis of Hasenstab et al. (2015) was able to connect to learning trends across the two diagnostic groups based on condition differentiation (in the P3 amplitudes), the analysis could not connect results to specific underlying cognitive processes. Note that in the time-frequency joint analysis, we are able to utilize time-frequency specific information along with scalp section, to connect results to specific processes. More specifically theta power contributions to P3 in the posterior scalp section are typically associated with sensory and visual processing, while theta power contributions to P3 in the frontal scalp section are typically associated with orienting to novel stimulus by Bernat et al. (2007). Hence the time-frequency joint analysis brings an additional level of specificity to interpretations which is lacking in time domain only or frequency domain only analysis.

5. Discussion

We have proposed LTFT-ERP to model the longitudinal trends in the ERP signal over trials/ duration of an EEG experiment. Longitudinal changes during the course of an experiment may be the main interest in some studies exemplified by the implicit learning paradigm, where longitudinal trends capture speed and nature of learning among TD and ASD children. LTFT-ERP utilizes time-frequency transformations to retain valuable information from both the time and frequency domains, enhancing interpretability of the findings. In addition, under the assumption of constant direction of variation in the functional domain, the algorithm borrows information across trials (i.e. the longitudinal dimension) in identifying the leading eigenvectors in the functional domain (carrying both time and frequency information) through the proposed MDPCA. This is a major stabilization tool for the proposed algorithm as it enhances the signal and enables further modeling of the longitudinal trends in the longitudinal MDPCA scores, obtained through projections of the signal onto the leading functional eigenvectors. Application to the implicit learning paradigm uncovers distinct learning patterns throughout the experiment among children diagnosed with ASD and their typically developing peers.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

References

Bernat EM, Williams WJ, Gehring WJ, 2005. Decomposing ERP time-frequency energy using PCA. Clin. Neurophysiol. 116 (6), 1314–1334. 10.1016/j.clinph.2005.01.019. [PubMed: 15978494]

Bernat EM, Malone SM, Williams WJ, Patrick CJ, Iacono WG, 2007. Decomposing delta, theta, and alpha time-frequency ERP activity from a visual oddball task using PCA. Int. J. Psychophysiol. 64 (1), 62–74. 10.1016/j.ijpsycho.2006.07.015. [PubMed: 17027110]

Bugli C, Lambert P, 2006. Functional ANOVA with random functional effects: an application to event-related potentials modelling for electroencephalograms analysis. Stat. Med. 25 (21), 3718– 3739. 10.1002/sim.2464. [PubMed: 16372388]

Chen K, Müller HG, 2012. Modeling repeated functional observations. J. Am. Stat. Assoc. 107 (500), 1599–1609. 10.1080/01621459.2012.734196.

Chen K, Delicado P, Müller HG, 2017. Modelling function-valued stochastic processes, with applications to fertility dynamics. J. R. Stat. Soc., Ser. B, Stat. Methodol. 79 (1), 177–196. 10.1111/ rssb.12160.

Delorme A, Makeig S, 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J. Neurosci. Methods 134 (1), 9–21. 10.1016/ j.jneumeth.2003.10.009. [PubMed: 15102499]

Di CZ, Crainiceanu CM, Caffo BS, Punjabi NM, 2009. Multilevel functional principal component analysis. Ann. Appl. Stat 3 (1), 458–488. 10.1214/08-AOAS206. [PubMed: 20221415]

Dong J, Estes JP, Li G, entürk D, 2016. A two-step estimation approach for logistic varying coefficient modeling of longitudinal data. J. Stat. Plan. Inference 174, 38–51. 10.1016/ j.jspi.2016.01.012. [PubMed: 27103756]

Farge M, 1992. Wavelet transforms and their applications to turbulence. Annu. Rev. Fluid Mech. 24 (1), 395–458. 10.1146/annurev.fl.24.010192.002143.

Fiecas M, Ombao H, 2016. Modeling the evolution of dynamic brain processes during an associative learning experiment. J. Am. Stat. Assoc. 111 (516), 1440–1453. 10.1080/01621459.2016.1165683.

Gasser T, Molinari L, 1996. The analysis of the EEG. Stat. Methods Med. Res. 5 (1), 67–99. 10.1177/096228029600500105. [PubMed: 8743079]

Greven S, Crainiceanu C, Caffo B, Reich D, 2010. Longitudinal functional principal component analysis. Electron. J. Stat. 4, 1022–1054. 10.1214/10-EJS575. [PubMed: 21743825]

Harper J, Malone SM, Bernat EM, 2014. Theta and delta band activity explain N2 and P3 ERP component activity in a go/no-go task. Clin. Neurophysiol. 125 (1), 124–132. 10.1016/j.clinph.2013.06.025. [PubMed: 23891195]

Hasenstab K, Sugar CA, Telesca D, Mcevoy K, Jeste SS, entürk D, 2015. Identifying longitudinal trends within EEG experiments. Biometrics 71 (4), 1090–1100. 10.1111/biom.12347. [PubMed: 26195327]

Jeste SS, Kirkham N, entürk D, Hasenstab K, Sugar C, Kupelian C, Baker E, Sanders AJ, Shimizu C, Norona A, Paparella T, Freeman SF, Johnson SP, 2015. Electrophysiological evidence of heterogeneity in visual statistical learning in young children with ASD. Dev. Sci. 18 (1), 90–105. 10.1111/desc.12188. [PubMed: 24824992]

Krafty RT, Hall M, Guo W, 2011. Functional mixed effects spectral analysis. Biometrika 98 (3), 583–598. 10.1093/biomet/asr032. [PubMed: 26855437]

Krafty RT, Rosen O, Stoffer DS, Buysse DJ, Hall MH, 2017. Conditional spectral analysis of replicated multiple time series with application to nocturnal physiology. J. Am. Stat. Assoc. 112 (520), 1405–1416. 10.1080/01621459.2017.1281811. arXiv:1502.03153 [PubMed: 29430069]

Motta G, Ombao H, 2012. Evolutionary factor analysis of replicated time series. Biometrics 68 (3), 825–836. 10.1111/j.1541-0420.2012.01744.x. [PubMed: 22364516]

Ombao H, Ho MHR, 2006. Time-dependent frequency domain principal components analysis of multichannel non-stationary signals. Comput. Stat. Data Anal 50 (9), 2339–2360. 10.1016/j.csda.2004.12.011.

Park SY, Staicu AM, 2015. Longitudinal functional data analysis. Stat 4 (1), 212–226. 10.1002/sta4.89. arXiv:1506.08796. [PubMed: 26594358]

Rice JA, Wu CO, 2001. Nonparametric mixed effects models for unequally sampled noisy curves. Biometrics 57 (1), 253–259. 10.1111/j.0006-341X.2001.00253.x. [PubMed: 11252607]

Shi M, Weiss RE, Taylor JMG, 1996. An analysis of paediatric CD4 counts for acquired immune deficiency syndrome using flexible random curves. J. R. Stat. Soc., Ser. C, Appl. Stat. 45 (2), 151–163. 10.2307/2986151.

Tierney AL, Gabard-Durnam L, Vogel-Farley V, Tager-Flusberg H, Nelson CA, 2012. Developmental trajectories of resting EEG power: an endophenotype of autism spectrum disorder. PLoS ONE 7 (6), e39127. 10.1371/journal.pone.0039127. [PubMed: 22745707]

Torrence C, Compo GP, 1998. A practical guide to wavelet analysis. Bull. Am. Meteorol. Soc 79 (1), 61–78. 10.1175/1520-0477(1998)079<0061:APGTWA>2.0.CO;2.
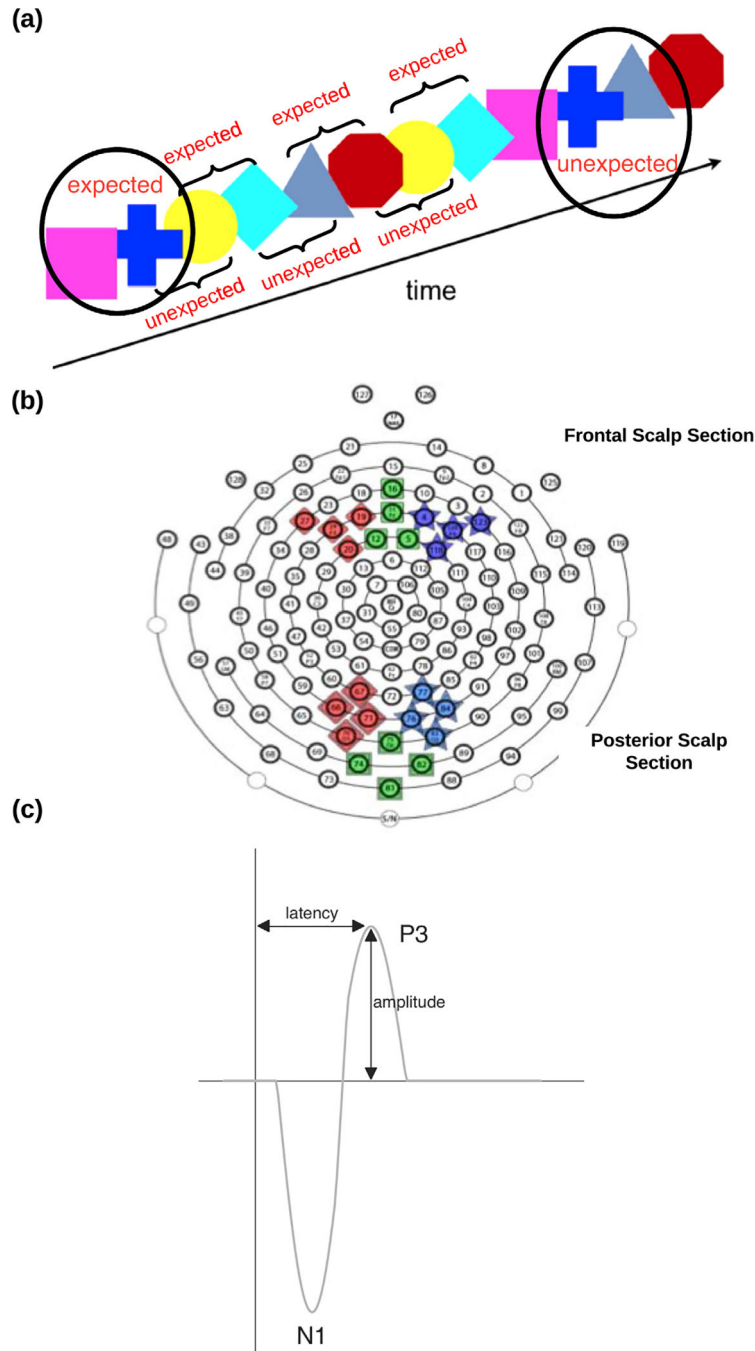
Fig. 1.
(a) Visualization of the implicit learning paradigm. The continuous stream of six-colored shapes are organized into three familiar pairs. The "expected" condition is defined as the transition between shapes within a shape pair, and the "unexpected" condition is defined as the transition between shape pairs. (b) The 24 electrodes of interest analyzed in the implicit learning paradigm in six total scalp regions (each containing four electrodes) within two scalp sections (frontal and posterior). (c) A depiction of the ERP phasic components P3 and N1 in the implicit learning paradigm.
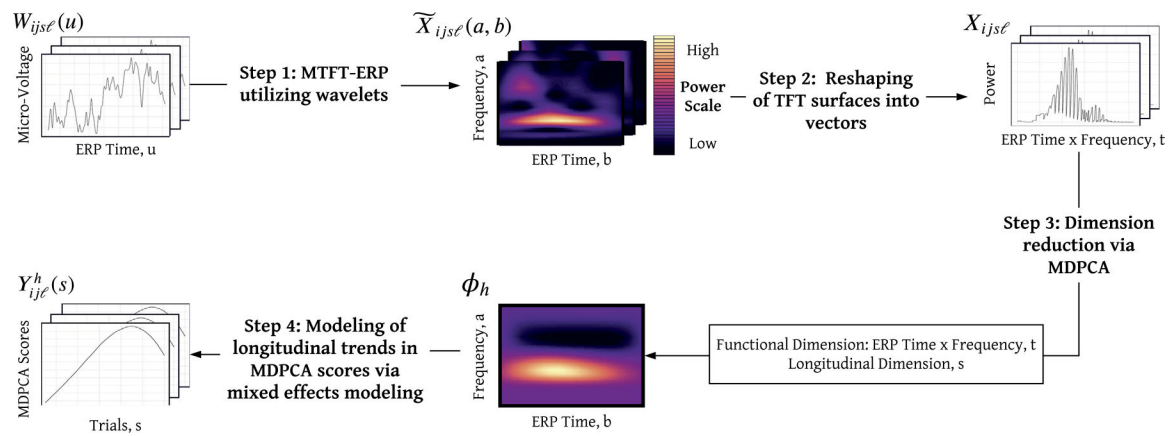
Fig. 2.
A flowchart of the LTFT-ERP algorithm. For each subject i, electrode j, on trial s and condition ℓ Step 1 transforms the ERP waveform $W_{ijj\ell}(u)$ into the TFT power surface $\widetilde{X}_{ijs\ell}(a,b)$ using the wavelet transformation. Step 2 reshapes the TFT power surface $\widetilde{X}_{ijs\ell}(a,b)$ into a vector $x_{ijs\ell}$ in t where t denotes the functional dimension of ERP time × frequency. The longitudinal dimension is trials s, where Step 3 performs dimension reduction via MDPCA to target the hth leading eigenvector φ_{h} in the functional dimension. The final step models the MDPCA scores as a function of trials (s) via mixed effects modeling.
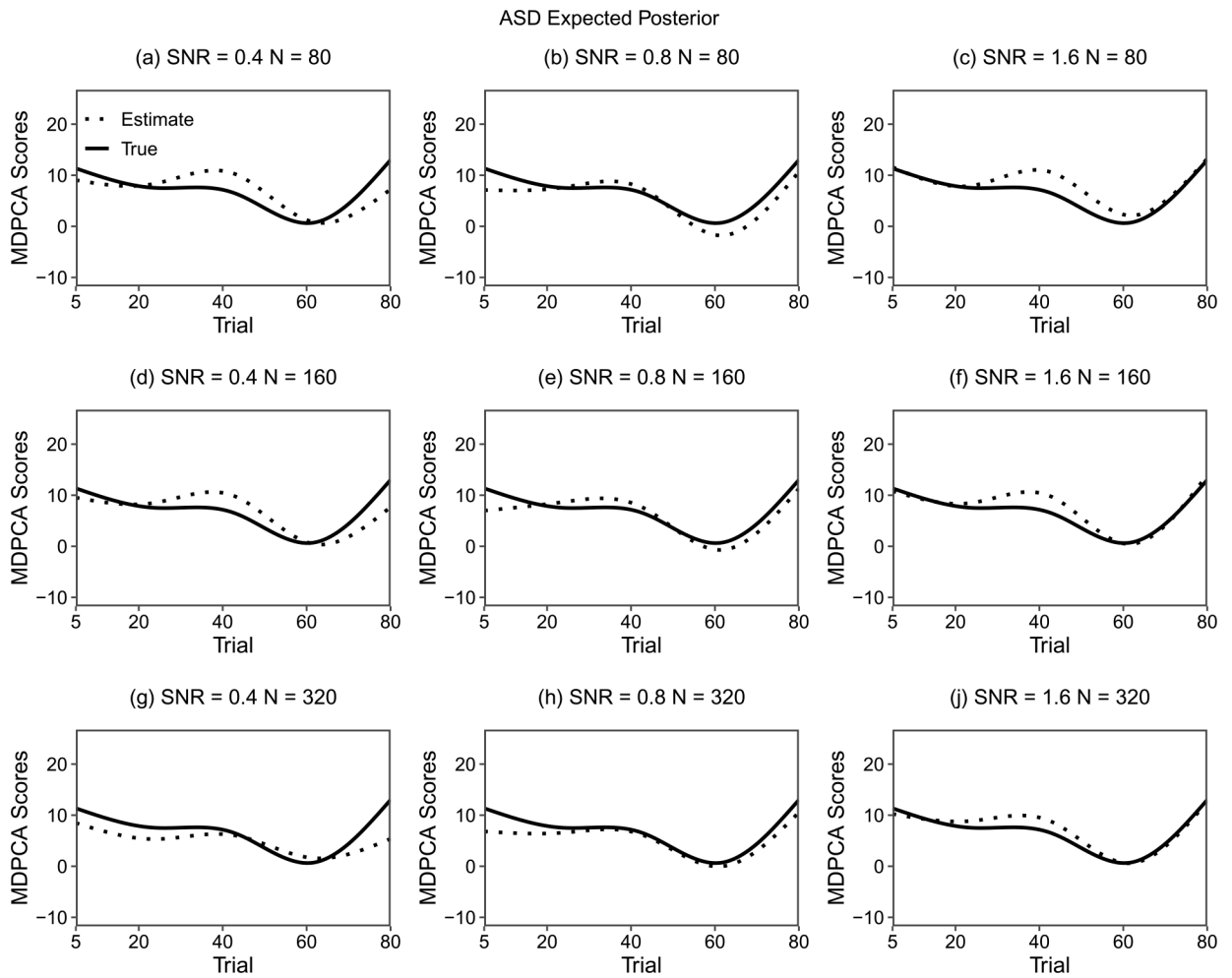
ASD Expected Posterior

Fig. 3.
True (solid) and estimated (dotted) fixed effects mean trajectories from the run with the median ME value for the MDPCA scores of the subgroup ASD expected posterior, at varying SNRs (rows) and sample sizes (columns).
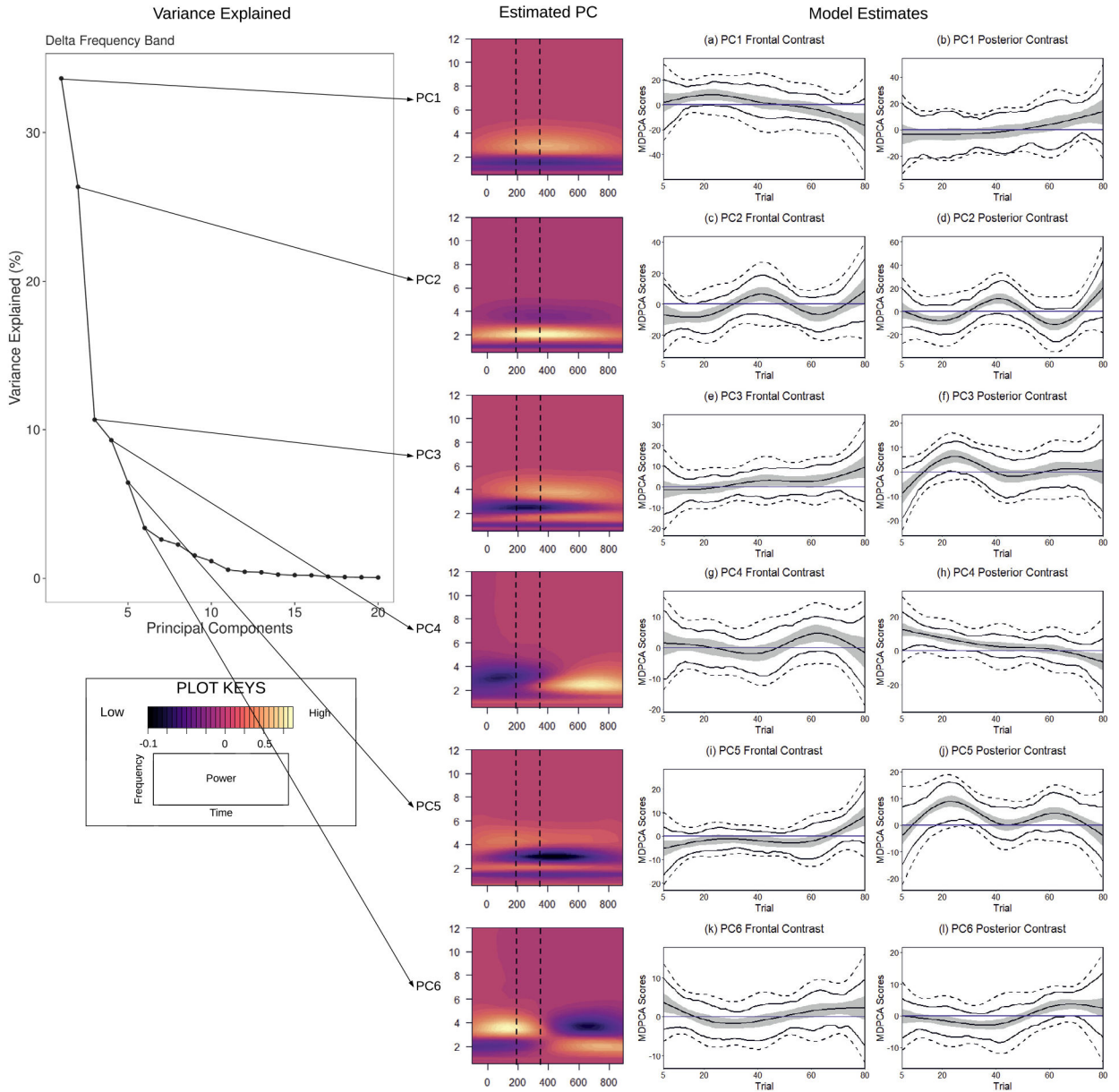
Fig. 4.
Summary of results from the proposed LTFT-ERP algorithm from the delta frequency band. The six estimated leading functional eigenvectors and their corresponding percent of variance explained are depicted. Contrasts for ((ASD expected - ASD unexpected) - (TD expected - TD unexpected)), based on the mixed effects modeling of the MDPCA scores, are also depicted. The contrasts and the associated 95% pointwise and simultaneous bootstrap intervals, based on resampling from subjects with replacement, are given in solid black and dashed black, respectively, while the 95% pointwise confidence intervals based on the mixed effects modeling are shaded in gray. A blue horizontal line at zero is included for ease of interpretation.
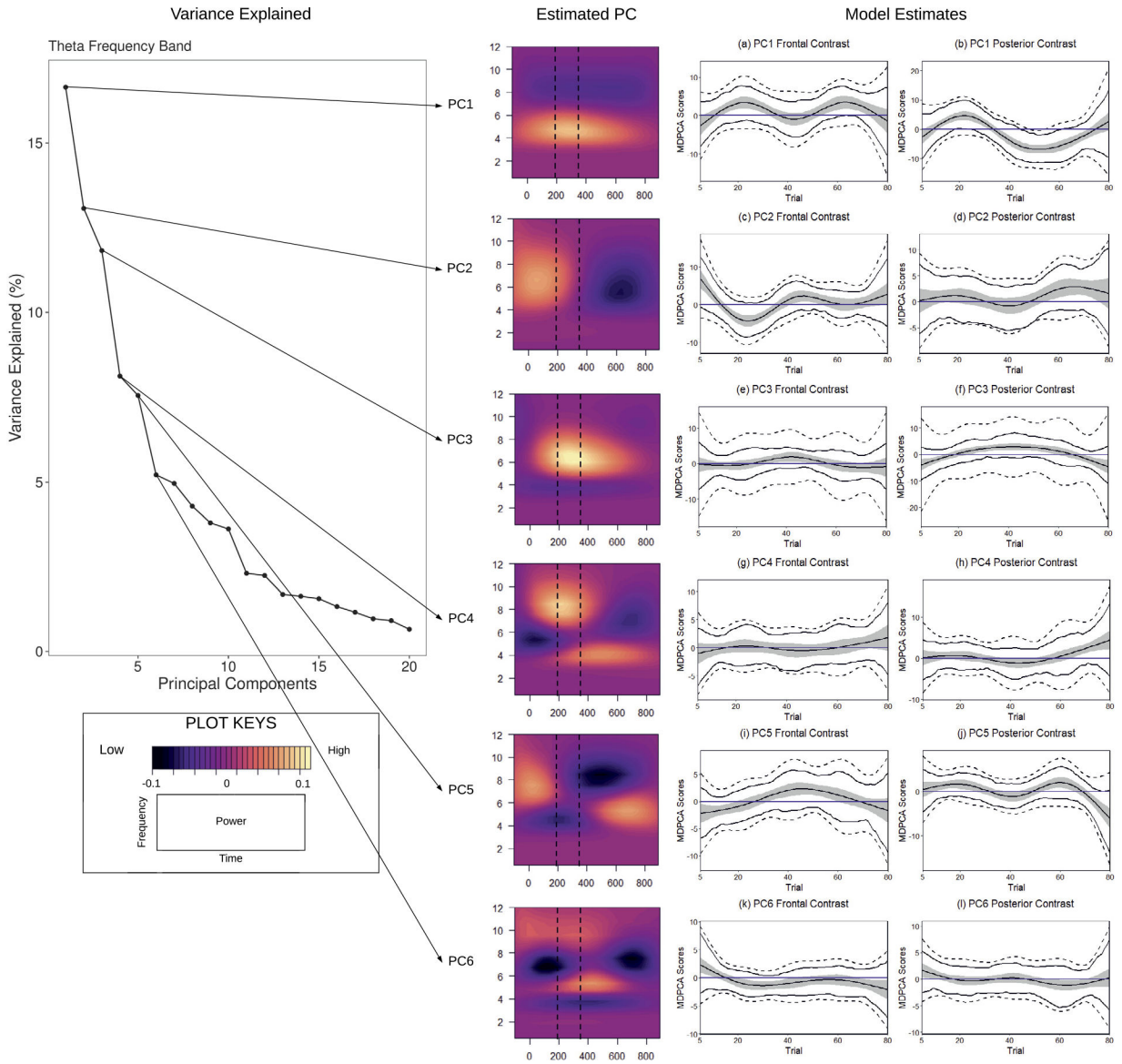
Fig. 5.
Summary of results from the proposed LTFT-ERP algorithm from the theta frequency band. The six estimated leading functional eigenvectors and their corresponding percent of variance explained are depicted. Contrasts for ((ASD expected - ASD unexpected) - (TD expected - TD unexpected)), based on the mixed effects modeling of the MDPCA scores, are also depicted. The contrasts and the associated 95% pointwise and simultaneous bootstrap intervals, based on resampling from subjects with replacement, are given in solid black and dashed black, respectively, while the 95% pointwise confidence intervals based on the mixed effects modeling are shaded in gray. A blue horizontal line at zero is included for ease of interpretation.
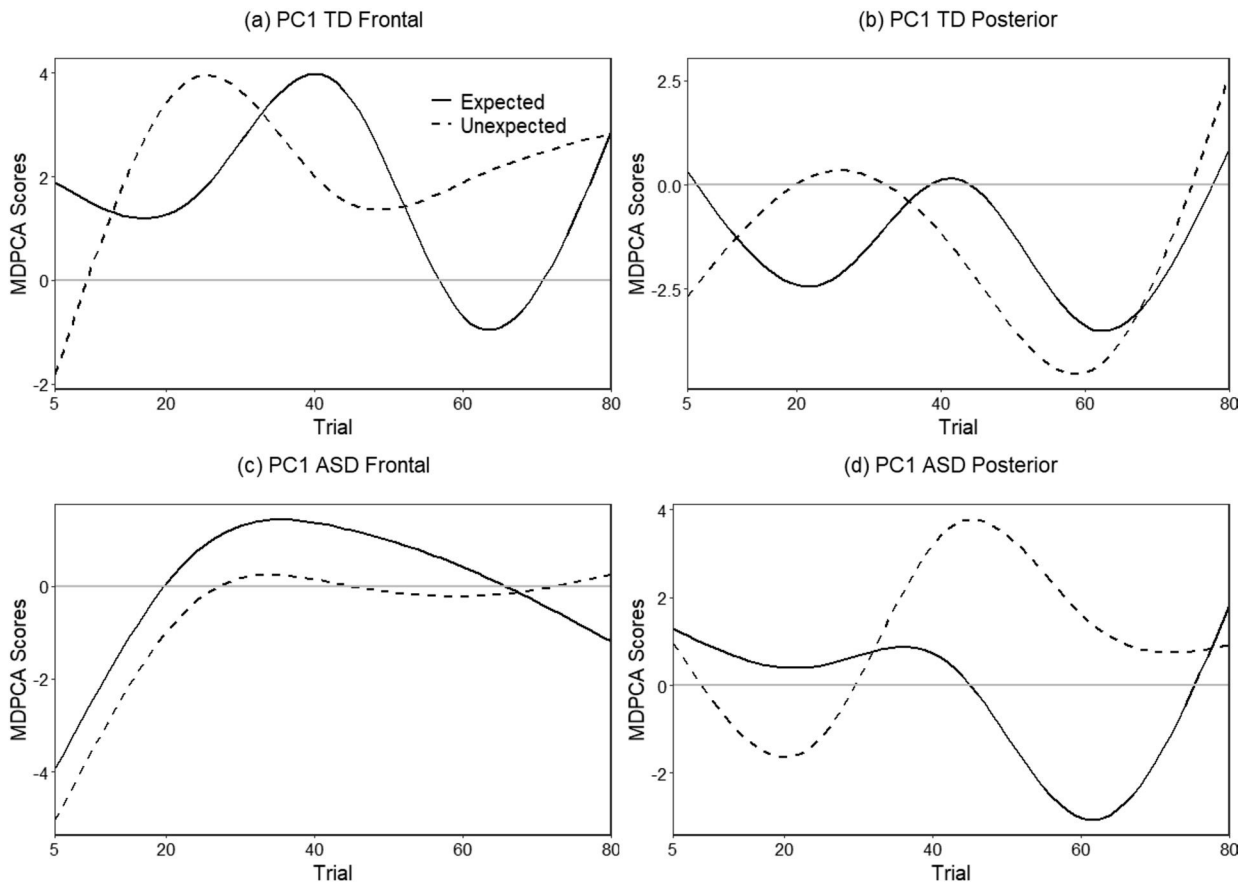
Fig. 6.
Estimated mean trajectories of the MDPCA scores for the leading first eigencomponent in the theta frequency band for each group, scalp section, and condition (expected solid and unexpected dashed). A gray horizontal line at zero is included for ease of interpretation.

Table 1

Medians and (10th, 90th) percentiles of simulation performance metrics (ME and PE) from 200 Monte Carlo runs at varying SNRs (SNR = 0.4, 0.8, and 1.6) and sample sizes (N = 80, 160, and 320).

N	SNR	Mean Power	Eigenvector	MDPCA Scores	
		ME	ME	ME	PE
80	0.4	0.81460 (0.80966, 0.82040)	0.89373 (0.88256, 0.90372)	0.376 (0.324, 0.497)	0.564 (0.534, 0.621)
	0.8	0.40759 (0.40359, 0.41065)	0.69500 (0.65928, 0.72499)	0.293 (0.228, 0.415)	0.426 (0.380, 0.498)
	1.6	0.20421 (0.20144, 0.20652)	0.32873 (0.28398, 0.37164)	0.255 (0.184, 0.554)	0.272 (0.221, 0.728)
160	0.4	0.81453 (0.81069, 0.81836)	0.89293 (0.88269, 0.90094)	0.344 (0.309, 0.389)	0.561 (0.532, 0.590)
	0.8	0.40756 (0.40489, 0.41020)	0.69260 (0.67283, 0.71025)	0.245 (0.203, 0.345)	0.418 (0.388, 0.507)
	1.6	0.20394 (0.20169, 0.20564)	0.32145 (0.28761, 0.35842)	0.206 (0.146, 0.422)	0.265 (0.224, 0.510)
320	0.4	0.81462 (0.81215, 0.81717)	0.89206 (0.88639, 0.89865)	0.333 (0.300, 0.429)	0.566 (0.544, 0.670)
	0.8	0.40766 (0.40579, 0.40918)	0.68958 (0.67272, 0.70487)	0.236 (0.198, 0.356)	0.418 (0.394, 0.599)
	1.6	0.20382 (0.20255, 0.20522)	0.32116 (0.29401, 0.34458)	0.182 (0.140, 0.376)	0.270 (0.237, 0.573)

Table 2

Medians and (10th, 90th) percentiles of simulation performance metrics (ME and PE) from 200 Monte Carlo run at varying SNRs (SNR = 0.4, 0.8, and 1.6) and sample sizes (N = 80, 160, and 320) for each of the eight subgroups determined by diagnostic group, condition and scalp section.

ME of MDPCA Scores

Group	N	SNR	Frontal		Posterior	
			Expected	Unexpected	Expected	Unexpected
ASD	80	0.4	0.547 (0.281, 0.882)	0.567 (0.377, 0.876)	0.348 (0.252, 0.452)	0.235 (0.171, 0.395)
		0.8	0.387 (0.232, 0.769)	0.402 (0.241, 0.721)	0.265 (0.169, 0.375)	0.167 (0.105, 0.319)
		1.6	0.384 (0.195, 0.853)	0.382 (0.214, 0.835)	0.242 (0.126, 0.513)	0.202 (0.107, 0.505)
	160	0.4	0.488 (0.272, 0.738)	0.507 (0.332, 0.734)	0.314 (0.253, 0.401)	0.219 (0.160, 0.279)
		0.8	0.331 (0.214, 0.602)	0.352 (0.234, 0.579)	0.224 (0.151, 0.312)	0.130 (0.077, 0.237)
		1.6	0.242 (0.133, 0.664)	0.245 (0.132, 0.636)	0.189 (0.115, 0.367)	0.171 (0.093, 0.349)
	320	0.4	0.463 (0.293, 0.728)	0.501 (0.362, 0.691)	0.311 (0.242, 0.389)	0.209 (0.156, 0.301)
		0.8	0.345 (0.205, 0.566)	0.328 (0.224, 0.506)	0.205 (0.145, 0.293)	0.120 (0.080, 0.229)
		1.6	0.230 (0.108, 0.480)	0.228 (0.122, 0.459)	0.166 (0.103, 0.322)	0.155 (0.077, 0.296)
TD	80	0.4	0.266 (0.178, 0.465)	0.284 (0.162, 0.523)	0.632 (0.484, 0.798)	0.679 (0.422, 0.989)
		0.8	0.202 (0.107, 0.340)	0.235 (0.127, 0.439)	0.431 (0.279, 0.649)	0.653 (0.400, 0.949)
		1.6	0.142 (0.070, 0.564)	0.174 (0.098, 0.571)	0.330 (0.191, 0.726)	0.595 (0.345, 1.137)
	160	0.4	0.238 (0.175, 0.336)	0.258 (0.171, 0.382)	0.580 (0.445, 0.682)	0.594 (0.413, 0.805)
		0.8	0.178 (0.112, 0.300)	0.206 (0.134, 0.351)	0.377 (0.265, 0.542)	0.551 (0.402, 0.776)
		1.6	0.115 (0.063, 0.349)	0.129 (0.075, 0.405)	0.289 (0.153, 0.546)	0.514 (0.285, 0.844)
	320	0.4	0.237 (0.188, 0.388)	0.267 (0.187, 0.420)	0.571 (0.464, 0.683)	0.561 (0.402, 0.729)
		0.8	0.165 (0.115, 0.336)	0.206 (0.134, 0.367)	0.356 (0.266, 0.512)	0.526 (0.400, 0.710)
		1.6	0.093 (0.053, 0.393)	0.115 (0.072, 0.414)	0.266 (0.171, 0.495)	0.466 (0.309, 0.753)

PE of MDPCA Scores

Group	N	SNR	Frontal		Posterior	
			Expected	Unexpected	Expected	Unexpected
ASD	80	0.4	0.617 (0.564, 0.699)	0.625 (0.572, 0.709)	0.561 (0.503, 0.662)	0.519 (0.471, 0.618)
		0.8	0.464 (0.406, 0.541)	0.466 (0.407, 0.542)	0.418 (0.359, 0.517)	0.387 (0.327, 0.495)
		1.6	0.285 (0.228, 0.751)	0.290 (0.229, 0.751)	0.272 (0.204, 0.725)	0.257 (0.193, 0.682)
	160	0.4	0.607 (0.558, 0.664)	0.613 (0.570, 0.667)	0.551 (0.510, 0.602)	0.511 (0.472, 0.563)
		0.8	0.454 (0.416, 0.544)	0.456 (0.416, 0.539)	0.414 (0.374, 0.519)	0.381 (0.344, 0.486)
		1.6	0.283 (0.239, 0.540)	0.286 (0.239, 0.539)	0.258 (0.215, 0.557)	0.245 (0.199, 0.525)
	320	0.4	0.612 (0.580, 0.712)	0.619 (0.584, 0.711)	0.554 (0.528, 0.647)	0.516 (0.488, 0.606)
		0.8	0.455 (0.421, 0.653)	0.457 (0.426, 0.645)	0.410 (0.380, 0.606)	0.379 (0.352, 0.569)
		1.6	0.289 (0.254, 0.591)	0.291 (0.255, 0.584)	0.267 (0.225, 0.594)	0.256 (0.215, 0.560)
TD	80	0.4	0.505 (0.458, 0.584)	0.524 (0.477, 0.612)	0.615 (0.563, 0.715)	0.611 (0.556, 0.710)
		0.8	0.378 (0.326, 0.465)	0.398 (0.342, 0.476)	0.451 (0.390, 0.537)	0.466 (0.400, 0.550)
		1.6	0.235 (0.185, 0.667)	0.248 (0.198, 0.692)	0.298 (0.230, 0.734)	0.312 (0.238, 0.698)
	160	0.4	0.501 (0.458, 0.555)	0.522 (0.477, 0.569)	0.610 (0.570, 0.664)	0.608 (0.566, 0.653)

	0.8	0.375 (0.341, 0.442)	0.395 (0.358, 0.458)	0.453 (0.411, 0.559)	0.469 (0.421, 0.560)
	1.6	0.231 (0.196, 0.486)	0.244 (0.209, 0.504)	0.279 (0.234, 0.540)	0.296 (0.244, 0.558)
320	0.4	0.505 (0.473, 0.630)	0.525 (0.492, 0.648)	0.614 (0.583, 0.722)	0.613 (0.575, 0.718)
	0.8	0.375 (0.349, 0.544)	0.394 (0.364, 0.554)	0.450 (0.417, 0.595)	0.466 (0.429, 0.597)
	1.6	0.234 (0.206, 0.533)	0.250 (0.219, 0.540)	0.289 (0.253, 0.563)	0.306 (0.264, 0.571)