Covariate-adjusted hybrid principal components analysis for region-referenced functional EEG data

Aaron Wolfe Scheffler*, Abigail Dickinson, Charlotte DiStefano, Shafali Jeste, and Damla Şentürk

Electroencephalography (EEG) studies produce region-referenced functional data via EEG signals recorded across scalp electrodes. The high-dimensional data can be used to contrast neurodevelopmental trajectories between diagnostic groups, for example between typically developing (TD) children and children with autism spectrum disorder (ASD). Valid inference requires characterization of the complex EEG dependency structure as well as covariate-dependent heteroscedasticity, such as changes in variation over developmental age. In our motivating study, EEG data is collected on TD and ASD children aged two to twelve years old. The peak alpha frequency, a prominent peak in the alpha spectrum, is a biomarker linked to neurodevelopment that shifts as children age. To retain information, we model patterns of alpha spectral variation, rather than just the peak location, regionally across the scalp and chronologically across development. We propose a covariate-adjusted hybrid principal components analysis (CA-HPCA) for EEG data, which utilizes both vector and functional principal components analysis while simultaneously adjusting for covariate-dependent heteroscedasticity. CA-HPCA assumes the covariance process is weakly separable conditional on observed covariates, allowing for covariate-adjustments to be made on the marginal covariances rather than the full covariance leading to stable and computationally efficient estimation. The proposed methodology provides novel insights into neurodevelopmental differences between TD and ASD children.

KEYWORDS AND PHRASES: Autism spectrum disorder, Covariate-adjustments, Electroencephalography, Functional data analysis, Heteroscedasticity.

1. INTRODUCTION

Despite the numerous developmental delays observed in children with autism spectrum disorder (ASD) compared to their typically developing peers (TD), the neural mechanisms underpinning these delays are not well characterized. To address this gap, our motivating study collected resting-state electroencephalograms (EEG) on TD and ASD children aged two to twelve years old, making it possible to

*Corresponding author.

contrast neural processes between the two diagnostic groups over a wide developmental range. EEG and magnetoencephalography (MEG) characterize cortical and intracortical brain activity, respectively, via the measurement of electrical potentials and their corresponding oscillatory dynamics (i.e. spectral characteristics). Recent studies in cognitive development using both EEG and MEG highlight the peak alpha frequency (PAF), defined as the location of a single prominent peak in the spectral density within the alpha frequency band (6–14 Hz), as a potential biomarker associated with autism diagnosis [16, 14, 15]. Specifically, the location of the PAF tends to shift from lower to higher frequencies as TD children age but this chronological shift is notably delayed or absent in ASD children [38, 30, 14, 15]. This trend is observed in our motivating data from a temporal electrode (T8) where the PAF, identifiable as 'humps' in age-specific slices of the group-specific bivariate mean alpha spectral density (across age and frequency), increases in frequency with age for TD children but not for ASD children (Figure 1 (a)).

Although the PAF holds promise as a biomarker for neural development in TD and ASD children, emphasis on the identification of a single peak produces considerable drawbacks. Not only is estimation of subject-specific PAF error prone due to the presence of noise and multiple local maxima [11] but also identification of a single peak frequency inherently reduces the alpha spectral band to a single scalar summary resulting in a loss of information. To avoid these limitations, we follow Scheffler et al. [36] and consider the entire spectral density across the alpha band as a functional measurement of neural activity.

In our motivating data, EEG signals are recorded from a high-density electrode array for several minutes and the time series at each channel is divided into overlapping two-second segments prior to Fast Fourier Transform (FFT) to the frequency domain. Spectral information is averaged across segments to boost the signal-to-noise ratio and the resulting data form region-referenced functional data with electrodes and spectral densities referred to as the regional and functional dimensions, respectively.

We focus on modeling and contrasting patterns of alpha spectral variation regionally across the scalp and chronologically across development for both the ASD and TD diagnostic groups. Previous research clearly shows that alpha spectral dynamics differ as a function of age between
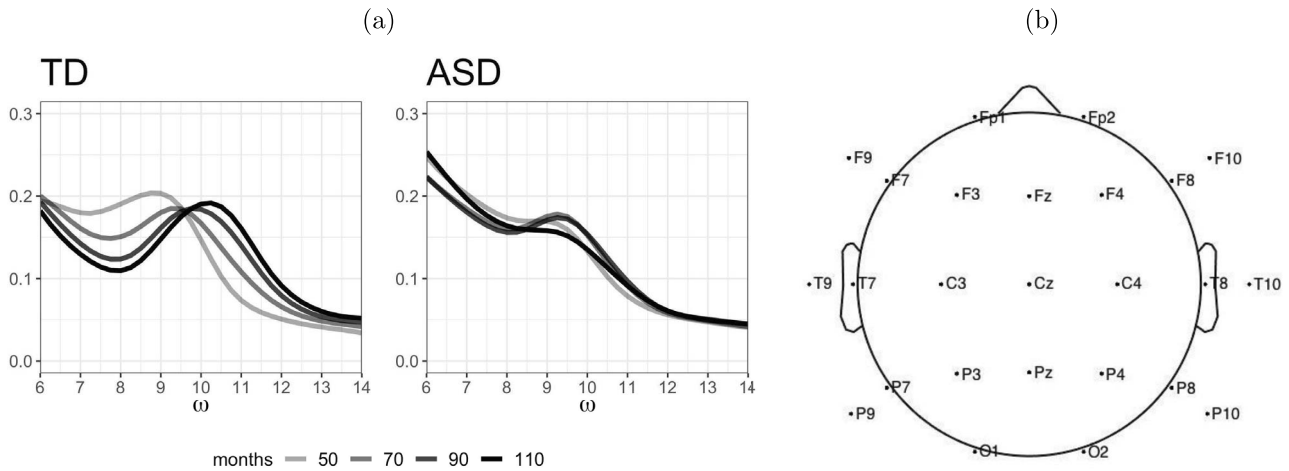
Figure 1. (a) Slices of the group-specific bivariate mean alpha spectral density (across age and frequency (6–14 Hz)) at ages 50, 70, 90 and 110 months from the T8 electrode. Darker lines correspond to older children. (b) A schematic diagram of the 10–20 25 electrode montage observed in the EEG data.

TD and ASD children and to assume a constant covariance structure across development risks missing important findings. To preserve developmental information, we propose a covariate-adjusted hybrid principal components analysis (CA-HPCA) that models variation in region-referenced functional data while simultaneously allowing the patterns of variation to change as a function of subject-specific covariates. CA-HPCA assumes the covariance process is weakly separable conditional on observed covariates, allowing for covariate-adjustments to be made on the marginal covariances rather than the full covariance leading to stable and computationally efficient estimation.

Since the introduction of the functional principal components analysis (FPCA) expansion (i.e. Karhunen–Loève expansion; [24, 28]), a detailed literature has developed around the estimation of functional principal scores and components for both densely and sparsely observed functional data along a single dimension (see Wang et al. [39] for a thorough review). In recent years, the literature surrounding FPCA has shifted to consider functional data with more complex dependency structures, including repeatedly measured functional data [12, 13, 26, 32, 31, 45], longitudinally observed functional data [19, 8, 33, 7, 29], spatially correlated functional data [2, 18, 44, 37, 27], both spatially and longitudinally observed functional data [21, 35], and multivariate functional data [22, 9, 20]. While these methods permit modeling of high dimensional functional covariances, they are unable to adjust for covariates in the analysis of higher dimensional functional data where the covariate may introduce heteroscedasticity to the functional dependency structure (e.g. due to chronological age).

In the simplified context of one-dimensional functional data, existing methods allow for covariate-adjustments to both the functional mean and covariance. Generally, the functional mean is smoothed across the covariate-domain, or calculated for each class in the case of discrete covariates. Covariate-adjustments are made to the functional covariance in two ways: either both the eigenvalues and eigenfunctions of the functional covariance are allowed to change as a function of observed covariates, or the eigenfunctions are assumed to be constant across the covariate dimension but their corresponding eigenvalues (as well as principal scores) are covariate-dependent. In the former class, Cardot [5] proposed a non-parametric covariate-adjusted FPCA in the context of dense functional data and Jiang and Wang [23] extended covariate-adjusted FPCA to noisy or sparse settings by estimating subject-specific scores using conditional expectation. In both cases, covariance estimation is performed non-parametrically by simultaneous smoothing across the covariate and functional domains via kernel methods. By fixing eigenfunctions across the covariate domain, Chiou et al. [10] introduced a semi-parametric functional regression model that estimates covariate-dependent principal scores using a single-index model and Backenroth et al. [1] developed a heteroscedastic FPCA for repeatedly measured curves that models eigenvalues as an exponential function of covariate and subject-dependent effects. We note that a parallel but distinct time series literature exists which focuses on estimation of covariate modulated spectral densities [17, 25, 4] but these works primarily focus on directly modeling non-stationary spectra as opposed functional data generally and do not embed covariate-adjustments into simplifying assumptions of the high-dimensional covariance structure (i.e. weak separability).

Our proposed covariate-adjusted hybrid principal components analysis (CA-HPCA) combines existing one-dimensional methods for covariate-dependent func-

tional heteroscedasticity with recent advances in multi-dimensional FPCA to allow covariate-adjustments in the context of region-referenced functional data. We briefly explore the methodological contributions of our proposed model and the resulting computational gains. A central theme in FPCA decompositions for multi-dimensional functional data is the use of simplifying assumptions regarding the covariance structure to ease estimation. A flexible approach in modeling two-dimensional functional data is to assume weak separability of the covariance process [7, 29] in which the marginal covariances along each dimension are targeted and the full covariance is projected onto a tensor basis formed from the corresponding marginal eigenfunctions. Thus, estimation is reduced from that of the total covariance in four-dimensions to the marginal covariances in two-dimensions for which efficient two-dimensional smoothers exist. Scheffler et al. [35] extended weak separability to region-referenced functional EEG data by proposing a hybrid principal components analysis (HPCA) that includes a discrete regional dimension but this model does not allow for the mean or covariance to change across development as needed in our application. We leverage the simplifying assumptions and computational efficiency of HPCA by introducing covariate-dependence to the functional mean and covariance which allows the marginal eigenvalues and eigenfunctions to change across the covariate domain.

In addition to the flexible modeling framework, CA-HPCA also introduces major computational savings. These savings are related to the the addition of a covariate dimension to estimation of the marginal covariances which for a scalar covariate requires smoothing across three dimensions. Previous methods such as Cardot [5] and Jiang and Wang [23] utilized kernel methods to estimate covariate-dependent marginal covariances but these approaches are computationally intensive and scale poorly with the introduction of additional covariates. To address this challenge, we extend the fast functional covariance smoothing proposed by Cederbaum et al. [6] to allow for covariate-adjustments by including an additional basis along the covariate dimension. Thus, CA-HPCA generalizes covariate-adjustments to high-dimensional functional covariances and achieves substantial reduction in computational burden by applying covariate adjustments to the marginal covariances with subsequent estimation performed via cutting-edge fast covariance smoothers.

A mixed effects framework is proposed to estimate the subject-specific scores with variance components that are a function of observed covariates. The estimated model components can be coupled with a parametric bootstrap resampling procedure to allow inference in the form of hypothesis testing and point-wise confidence intervals. We apply the proposed procedure to assess differences in alpha spectral dynamics between the TD and ASD groups across development. The remaining sections are organized as follows. Sec-tion 2 introduces the proposed CA-HPCA and Section 3 describes the corresponding estimation procedure. Application of the proposed method to our motivating EEG data follows in Section 4. Section 5 studies the finite-sample properties of CA-HPCA via extensive simulations. Section 6 concludes with a brief summary and discussion.

2. COVARIATE-ADJUSTED HYBRID PRINCIPAL COMPONENTS ANALYSIS (CA-HPCA)

Consider a region-referenced functional process observed in the presence of some continuous non-functional covariate $a_i \in \mathcal{A}$, $Y_{di}(a_i, r, \omega)$, for subject i, $i = 1, \ldots, n_d$, from group d, $d = 1, \ldots, D$, in region r, $r = 1, \ldots, R$, and at frequency ω, $\omega \in \Omega$. We decompose $Y_{di}(a_i, r, \omega)$ additively such that the expectation and covariance of the process both depend on the covariate a_i,

$$Y_{di}(a_i, r, \omega) = \eta_d(a_i, r, \omega) + Z_{di}(a_i, r, \omega) + \epsilon_{di}(a_i, r, \omega),$$

where $\eta_d(a_i, r, \omega) = E\{Y_{di}(a_i, r, \omega)|a_i\}$ denotes the group-region mean function, $Z_{di}(a_i, r, \omega)$ denotes a mean zero region-referenced stochastic process with total variance $\boldsymbol{\Sigma}_{d,T}(a_i; r, \omega; r', \omega') = \text{cov}\{Z_{di}(a_i, r, \omega), Z_{di}(a_i, r', \omega')|a_i\}$, and $\epsilon_{di}(a_i, r, \omega)$ denotes measurement error with mean zero and variance σ_d^2 that is independent across the regional, functional, and covariate domains. We assume the group-region mean functions $\eta_d(a, r, \omega)$ are smooth in both the functional domain Ω and the non-functional domain $\mathcal{A}$ though we place no restrictions across the regional domain $\mathbb{R}^R$ which in EEG data can lack the ordering provided by continuity.

In the proposed CA-HPCA model, we assume that the total covariance $\boldsymbol{\Sigma}_{d,T}(a; r, \omega; r', \omega')$ is weakly separable for each $a \in \mathcal{A}$. Weak separability, a concept recently proposed by Lynch and Chen [29] for two-dimensional functional data and adapted by Scheffler et al. [35] to region-referenced functional EEG data, implies that a covariance can be expressed as a weighted sum of separable covariance components and that the direction of variation (i.e. eigenvectors or eigenfunctions) along one dimensions of the EEG data is the same across fixed slices of the other dimension. Specifically, this assumption allows a multi-dimensional functional process to be decomposed parsimoniously via tensors of one-dimensional eigencomponents obtained from marginal covariances along each dimension. Note that weak separability is more flexible than strong separability (i.e. separability) commonly utilized in spatiotemporal modeling which requires the total covariance, not just the directions of variation, is the same up to a constant for fixed slices of the other dimensions. Unlike previous applications of weak separability, we introduce covariate-dependent heteroscedasticity by assuming the total covariance is weakly separable conditional on observed covariates and that marginal covariances

in each dimension vary smoothly along the covariate domain. Let the covariate-dependent regional and functional marginal covariances be defined as,

$$
\begin{aligned}
\{\mathbf{\Sigma}_{d,\mathcal{R}}(a)\}_{r,r'} &= \int_{\Omega} \mathrm{cov}\{Z_{di}(a_i,r,\omega), Z_{di}(a_i,r',\omega)\}d\omega \\
&= \sum_{k=1}^{R} \tau_{dk,\mathcal{R}}(a_i)\mathrm{v}_{dk}(a_i,r)\mathrm{v}_{dk}(a_i,r'), \\
\mathbf{\Sigma}_{d,\Omega}(a,\omega,\omega') &= \sum_{r=1}^{R} \mathrm{cov}\{Z_{di}(a_i,r,\omega), Z_{di}(a_i,r,\omega')\} \\
&= \sum_{\ell=1}^{\infty} \tau_{d\ell,\Omega}(a_i)\phi_{d\ell}(a_i,\omega)\phi_{d\ell}(a_i,\omega'),
\end{aligned}
$$

where $\mathrm{v}_{dk}(a,r)$ are the covariate-dependent eigenvectors of the regional marginal covariance matrix $\{\mathbf{\Sigma}_{d,\mathcal{R}}(a)\}_{r,r'}$, $\phi_{d\ell}(a,\omega)$ are the covariate-dependent eigenfunctions of the functional marginal covariance surface $\mathbf{\Sigma}_{d,\Omega}(a,\omega,\omega')$, and $\tau_{dk,\mathcal{R}}(a)$ and $\tau_{d\ell,\Omega}(a)$ are the regional and functional covariate-dependent marginal eigenvalues, respectively. Thus, there exists an orthogonal expansion of the covariate-dependent marginal covariances in terms of covariate-dependent marginal eigenvectors and eigenfunctions.

Utilizing the covariate-dependent eigenvectors and eigenfunctions, the covariate-adjusted hybrid principal components decomposition (**CA-HPCA**) of $Y_{di}(a_i,r,\omega)$ is given as,

$$
\begin{aligned}
Y_{di}(a_i,r,\omega) &= \eta_d(a_i,r,\omega) + Z_{di}(a_i,r,\omega) + \epsilon_{di}(a_i,r,\omega) \\
&= \eta_d(a_i,r,\omega) + \sum_{k=1}^{R}\sum_{\ell=1}^{\infty} \xi_{di,k\ell}(a_i)\mathrm{v}_{dk}(a_i,r) \\
&\times \phi_{d\ell}(a_i,\omega) + \epsilon_{di}(a_i,r,\omega),
\end{aligned}
$$

where $\xi_{di,k\ell}(a_i)$ are uncorrelated subject-specific scores defined through the projection of the region-referenced stochastic process onto the covariate-dependent tensor basis, $\langle Z_{di}(a_i,r,\omega), \mathrm{v}_{dk}(a_i,r)\phi_{d\ell}(a_i,\omega)\rangle = \sum_{r=1}^{R}\int Z_{di}(a_i,r,\omega)\mathrm{v}_{dk}(a_i,r)\phi_{d\ell}(a_i,\omega)d\omega$. Note that under the assumption of weak separability the subject-specific scores are uncorrelated over regions and frequencies. The CA-HPCA decomposition leads to the decomposition of the total covariance $\mathbf{\Sigma}_{d,T}(a;r,\omega;r',\omega')$ as follows,

$$
\begin{aligned}
\mathbf{\Sigma}_{d,T}(a;r,\omega;r',\omega') &= \mathrm{cov}\{Z_{di}(a,r,\omega), Z_{di}(a,r',\omega')|a\} \\
&+ \sigma_d^2\delta(a;r,\omega;r',\omega') \\
&= \sum_{k=1}^{R}\sum_{\ell=1}^{\infty} \tau_{d,k\ell}(a)\mathrm{v}_{dk}(a,r)\mathrm{v}_{dk}(a,r') \\
&\times\phi_{d\ell}(a,\omega)\phi_{d\ell}(a,\omega') \\
&+\sigma_d^2\delta(a;r,\omega;r',\omega'),
\end{aligned}
$$

where $\tau_{d,k\ell}(a) = \mathrm{var}\{\xi_{di,k\ell}(a)\}$ are covariate-dependent variance components and $\delta(a,r,\omega;a',r',\omega')$ denotes the indicator for $\{(a,r,\omega) = (a',r',\omega')\}$. Because the variance components $\tau_{d,k\ell}(a)$ are allowed to vary along the covariate domain, covariate-dependent heteroscedasticity is introduced not only through the covariate-dependent marginal eigencomponents but also in their relative contribution to the total variance.

In practice, the CA-HPCA decomposition is truncated to include only K_d and L_d covariate-dependent marginal eigencomponents for the regional and functional domains, respectively, with the number of components initially selected by the fraction of variance explained (FVE). One guideline is to include the minimum number of covariate-dependent marginal eigencomponents in the CA-HPCA expansion that explain at least 90% of variation in their respective covariate-dependent marginal covariances for each observed covariate value, though this may need to be relaxed in certain instances when the number of eigencomponents is excessively high for a few covariate values. The final number of components can be fixed after the subject-specific scores and their associated variance components are estimated via a mixed effects model proposed in Section 3.1 which allow enumeration of the total FVE in the observed data not just for the marginal covariances but the total covariance. Further details on the selection of the number of covariate-dependent marginal eigencomponents are presented in Section 3.1. As mentioned, the above model assumes that both the marginal directions of variation and their associated variance components for subject-specific scores are allowed to vary across the covariate domain. If instead we allow only the variance components to be covariate-dependent but restrict the marginal eigenfunctions and eigenvectors to be constant across the covariate-domain (i.e. the marginal directions of variation are common across the covariate domain), we produce a reduced model where the marginal covariance may be pooled across observed covariates. We defer specifics about this useful extension on the reduced CA-HPCA to Web Appendix A of the Supporting Information (http://intlpress.com/site/pub/files/_supp/sii/2022/0015/0002/SII-2022-0015-0002-s002.pdf). While the reduced CA-HPCA can lead to major computational savings, the assumption of covariate independent eigenfunctions may not be satisfied in every application. As an example, it was not found plausible in the context of our motivating EEG data where the directions of marginal variation were not constant across development.

3. ESTIMATION OF MODEL COMPONENTS AND INFERENCE

The following section details estimation of all CA-HPCA model components, including group-region mean functions, covariate-dependent marginal covariances, subject-specific decomposition scores and their associated variance components as well as procedures for inference made available

---

Algorithm 1 CA-HPCA Estimation Procedure

---

1. Estimation of group-region mean functions

   (a) Calculate $\hat{\eta}_d(a_i, r, \omega)$ by applying a bivariate penalized spline smoother to all observed data $\{a_i, \omega, Y_{di}(a_i, r, \omega) : i = 1, \ldots, n_d; a_i \in \mathcal{A}; \omega \in \Omega\}$.

   (b) Mean center each observation, $\widehat{Y}_{di}^c(a_i, r, \omega) = Y_{di}(a_i, r, \omega) - \hat{\eta}_d(a_i, r, \omega)$.

2. Estimation of covariate-dependent marginal covariances and measurement error variance

   (a) Calculate $\widehat{\boldsymbol{\Sigma}}_{d,\Omega}(a, \omega, \omega')$ and $\hat{\sigma}_{d,\Omega}^2$ by applying trivariate penalized spline smoothers to the products, $\{a_i, \omega, \omega', \widehat{Y}_{di}^c(a_i, r, \omega)\widehat{Y}_{di}^c(a_i, r, \omega') : i = 1, \ldots, n_d; a_i \in \mathcal{A}; r = 1, \ldots, R; \omega, \omega' \in \Omega\}$.

   (b) Calculate $\widehat{\boldsymbol{\Sigma}}_{d,\mathcal{R}}(a)$ by smoothing each (r, r') entry across $\mathcal{A}$. For $r \neq r'$, estimate $\{\widehat{\boldsymbol{\Sigma}}_{d,\mathcal{R}}(a)\}_{(r,r')}$ by applying a univariate kernel smoother to $\{a_i, r, r', \widehat{Y}_{di}^c(a_i, r, \omega)\widehat{Y}_{di}^c(a_i, r', \omega) : i = 1, \ldots, n_d; a_i \in \mathcal{A}; r \neq r' = 1, \ldots, R; \omega \in \Omega\}$. For $r = r'$, estimate $\{\widehat{\boldsymbol{\Sigma}}_{d,\mathcal{R}}(a)\}_{(r,r)}$ by applying a univariate kernel smoother to $\{a_i, r, r, \widehat{Y}_{di}^c(a_i, r, \omega)\widehat{Y}_{di}^c(a_i, r, \omega) - \hat{\sigma}_{d,\Omega}^2 : i = 1, \ldots, n_d; a_i \in \mathcal{A}; r = r' = 1, \ldots, R; \omega \in \Omega\}$.

3. Estimation of covariate-dependent marginal eigencomponents

   (a) For each unique value of a observed, employ FPCA on $\widehat{\boldsymbol{\Sigma}}_{d,\Omega}(a, \omega, \omega')$ to estimate the covariate-dependent eigenvalue, eigenfunction pairs, $\{\tau_{d\ell,\Omega}(a), \phi_{d\ell}(a, \omega) : \ell = 1, \ldots, L_d\}$.

   (b) For each unique value of a observed, employ PCA on $\widehat{\boldsymbol{\Sigma}}_{d,\mathcal{R}}(a)$ and to estimate the covariate-dependent eigenvalue, eigenvector pairs $\{\tau_{dk,\mathcal{R}}(a), \mathrm{v}_{dk}(a, r) : k = 1, \ldots K_d\}$.

4. Estimation of covariate-dependent variance components and subject-specific scores via linear mixed effects models

   (a) Calculate $\hat{\tau}_{dg}(a_i)$ and $\hat{\sigma}_d^2$ by fitting the proposed linear mixed effects model.

   (b) Select G_d' such that $FVE_{dG'} > .8$ for $d = 1, \ldots, D$.

   (c) Calculate $\hat{\xi}_{dig}(a_i)$ as the BLUP $\hat{\xi}_{dig}(a_i) = E\{\xi_{dig}(a_i)|\boldsymbol{Y}_{di}\}$ and form predictions $\widehat{Y}_{di}(a_i, r, \omega)$.

---

through the proposed linear mixed effects model. In addition, guidance is provided for the selection of the number of eigencomponents included in the proposed decomposition.

3.1 Estimation of CA-HPCA model components

We begin by introducing the CA-HPCA algorithm above and focus our discussion in this section more on the novel estimation procedures used in targeting covariate-dependent marginal covariances and variance components found in steps 2 and 4, respectively.

(1) Estimation of group-region mean functions: We calculate the estimated group-region mean function $\hat{\eta}_d(a_i, r, \omega)$ for each region via smoothing performed by projection onto a tensor basis formed by penalized marginal B-splines in the covariate and functional domains. Smoothing parameter selection is performed using restricted maximum likelihood (REML) methods. Assuming the estimated group-region mean functions lie in the space spanned by the marginal B-splines, the estimated group-region mean functions enjoy asymptotic consistency as discussed in [41].

(2) Estimation of covariate-dependent marginal covariances and measurement error variance: We estimate the covariate-dependent marginal covariances by assuming each two-dimensional marginal covariance varies smoothly over the covariate dimension. For the functional marginal covariance, $\boldsymbol{\Sigma}_{d,\Omega}(a, \omega, \omega')$, we extend the fast bivariate covariance

smoother of Cederbaum et al. [6] to include a third covariate dimension $a \in \mathcal{A}$. To briefly review, Cederbaum et al. [6] proposed a smooth method of moments approach to estimate covariance functions based on fast bivariate penalized splines. To achieve computational efficiency, their method leverages the symmetry of the covariance function to reduce the data used in estimation by targeting the upper triangle of the covariance surface (including the diagonal) and enforce symmetry constraints that reduce the number of spline coefficients needed for estimation.

We extend their approach via the development of a fast trivariate penalized spline smoother which incorporates covariate information through the introduction of a marginal spline basis along the covariate dimension. The resulting smoother maintains the computational efficiency of Cederbaum et al. [6] while simultaneously allowing the marginal functional covariance to vary smoothly along the covariate dimension. In the process of estimating the covariate-dependent marginal functional covariance, we obtain an initial estimate of the measurement error variance $\hat{\sigma}_{d,\Omega}^2$ as well by modeling the diagonal elements additively as a function of the marginal covariance and measurement error variance. Smoothing parameter selection is performed using REML methods. Note, the smooth method of moments estimator assumes independences of the cross products and homoscedastic Gaussian measurement error, common assumptions in the estimation of functional curves. Smooth covari-

ance estimators which allow for heteroscedasticity are explored in Xiao et al. [42] but are more computationally demanding. However, Cederbaum et al. [6] showed that estimates based on the working assumptions of independence and homoscedastic Gaussian measurement error are robust when these assumptions are violated and well worth the computational savings.

The regional marginal covariance $\{\mathbf{\Sigma}_{d,\mathcal{R}}(a)\}_{r,r'}$ is discrete in the regional dimension and thus not amenable to trivariate smoothers as the marginal functional covariance above. Therefore, we estimate the raw regional marginal covariance at each covariate-value by removing the measurement variance from the diagonals as in Scheffler et al. [35], and smooth the resulting matrices entry-by-entry along the covariate domain. To ensure a positive definite regional marginal covariance, we utilize a kernel function with a common bandwidth in smoothing each entry across the raw covariate-dependent regional marginal covariances. The optimal bandwidth is selected via leave-one-subject-out cross validation (LOSOCV). Our kernel smoother is the Nadaraya–Watson kernel weighted-average,

$$
\begin{aligned}
\{\widehat{\mathbf{\Sigma}}_{d,\mathcal{R}}(a_o)\}_{(r,r')} &= \Big\{ \sum_{i=1}^{n_d} \sum_{\omega \in \Omega} K_\lambda(a_i - a_o)\widehat{Y}_{di}^c(a_i,r,\omega) \\
&\times \; \widehat{Y}_{di}^c(a_i,r',\omega)\Big\} \Big/ \Big\{|\Omega| \sum_{i=1}^{n_d} K_\lambda(a_i - a_o)\Big\},
\end{aligned}
$$

where $K_\lambda(\cdot)$ is a kernel with bandwidth parameter λ, $\widehat{Y}_{di}^c(a_i,r,\omega) = Y_{di}(a_i,r,\omega) - \hat{\eta}_d(a_i,r,\omega)$ is the demeaned subject-level data, and $|\Omega|$ is the number of observed functional grid points. For example, in our data application and simulation study we make use of a Gaussian kernel function such that $K_\lambda(\cdot) = \exp\big(-\frac{(a_i - a_o)^2}{2\lambda}\big)$. The parameter λ is selected to minimize the $\mathrm{LOSOCV}(\lambda)$ statistic across all entries (r,r'),

$$
\begin{aligned}
\mathrm{LOSOCV}(\lambda) &= \sum_{r=1}^{R} \sum_{r'=1}^{r} \mathrm{LOSOCV}(\lambda,r,r'), \\
\mathrm{LOSOCV}(\lambda,r,r') &= \frac{1}{|\Omega|n_d} \sum_{i=1}^{n_d} \sum_{\omega \in \Omega} \big[\widehat{Y}_{di}^c(a_i,r,\omega) \\
&\times \; \widehat{Y}_{di}^c(a_i,r',\omega) - \{\widehat{\mathbf{\Sigma}}_{d,\mathcal{R}}^{(-i)}(a_i)\}_{(r,r')}\big]^2,
\end{aligned}
$$

where $\{\widehat{\mathbf{\Sigma}}_{d,\mathcal{R}}^{(-i)}(a_i)\}$ is the estimated smoothed marginal covariance matrix with the ith subject left out. As with any Nadaraya–Watson estimator, concerns can arise when the covariate space grows to dimensions higher than observed in our motivating data.

Thus, we introduce two novel covariate-dependent smoothers for the regional and functional marginal covariances that allow for calculation of the covariate-dependent marginal covariances that may be used for subsequent covariate-dependent eigendecompositions.

(3) Estimation of covariate-dependent marginal eigen-components: To estimate the covariate-dependent marginal eigencomponents we perform eigendecompositions at each fixed covariate-value as described in Scheffler et al. [35] retaining a common number of K_d and L_d covariate-dependent eigencomponents. We initially include K_d and L_d components that explain at least 90% of variation in their respective covariate-dependent marginal covariances for each observed covariate value.

(4) Estimation of covariate-dependent variance components and subject-specific scores via linear mixed effects models: We make use of the estimated group-region mean functions and covariate-dependent marginal eigencomponents to propose a linear mixed effects framework for estimation of the covariate-dependent variance components and measurement error variance. Under the assumption of joint normality of the covariate-dependent subject-specific scores and measurement error, the proposed mixed effects framework induces regularization and stability in modeling the data by enforcing a low-rank structure on the covariate-dependent variance components through projection of the corresponding precision components onto a smooth basis. The resulting variance components can be used to select the number of eigencomponents to include in the CA-HPCA decomposition by quantifying the proportion of variance explained, leading to parametric bootstrap based inference in the form of hypothesis testing and point-wise confidence intervals. We present the linear mixed effects modeling framework below.

To make our notation more compact, we replace the double index $k\ell$ in CA-HPCA decomposition truncated at K_d and L_d with a single index $g = (k-1) + K_d(\ell-1) + 1$,

$$
\begin{aligned}
Y_{di}(a_i,r,\omega) &= \eta_d(a_i,r,\omega) + \sum_{g=1}^{G_d} \xi_{dig}(a_i)\varphi_{dg}(a_i,r,\omega) \\
&+ \; \epsilon_{di}(a_i,r,\omega),
\end{aligned}
$$

where $\varphi_{dg}(a_i,r,\omega) = \mathrm{v}_{dk}(a_i,r)\phi_{d\ell}(a_i,\omega)$ is a covariate-dependent tensor basis formed from marginal eigencomponents, $\xi_{dig}(a_i) = \langle Z_{di}(a_i,r,\omega), \varphi_{dg}(a_i,r,\omega)\rangle$, $\tau_{dg}(a_i) = \mathrm{cov}\{\xi_{dig}(a_i)\}$, and $G_d = K_d L_d$. Let $\mathbf{Y}_{di}(a_i)$ represent the vectorized form of $Y_{di}(a_i,r,\omega)$ for subject i, $i = 1,\ldots,n_d$ observed along with covariate value a_i. Analogous vectorized forms for the group-region mean function, $\eta_d(a_i,r,\omega)$, the region-referenced stochastic process $Z_{di}(a_i,r,\omega)$, covariate-dependent tensor basis $\varphi_{dg}(a_i,r,\omega)$, and the the measurement error $\epsilon_{di}(a_i,r,\omega)$ are denoted by $\boldsymbol{\eta}_{di}(a_i)$, $\mathbf{Z}_{di}(a_i)$, $\boldsymbol{\varphi}_{dg}(a_i)$, and $\boldsymbol{\epsilon}_{di}(a_i)$, respectively. Under the assumption that $\boldsymbol{\xi}_{di}(a_i) = \{\xi_{di1}(a_i),\ldots,\xi_{diG}(a_i)\}$ and $\boldsymbol{\epsilon}_{di}(a_i)$ are jointly Gaussian and $\mathrm{cov}\{\boldsymbol{\xi}_{di}(a_i), \boldsymbol{\epsilon}_{di}(a_i)\} = \mathbf{0}$ at a fixed value of a_i, the proposed linear mixed effects model is given as

$$
\begin{aligned}
\mathbf{Y}_{di}(a_i) &= \boldsymbol{\eta}_{di}(a_i) + \mathbf{Z}_{di}(a_i) + \boldsymbol{\epsilon}_{di}(a_i) \\
&= \boldsymbol{\eta}_{di}(a_i) + \sum_{g=1}^{G_d} \xi_{dig}(a_i)\boldsymbol{\varphi}_{dg}(a_i) + \boldsymbol{\epsilon}_{di}(a_i),
\end{aligned}
\tag{1}
$$

for $i = 1, \ldots, n_d$. The model is fit separately in each group, $d = 1, \ldots, D$ and the regional and functional dependencies in $\boldsymbol{Y}_{di}(a_i)$ are induced through the subject-specific random effects $\xi_{dig}(a_i)$ in (1). Given the total covariance is weakly separable for fixed values of a, $\operatorname{cov}\{\xi_{dig}(a_i), \xi_{dig'}(a_i)\} = 0$ for $g \neq g'$ and thus the covariance matrix of the subject-specific scores possesses a diagonal diagonal structure, $\operatorname{cov}\{\boldsymbol{\xi}_{di}(a_i)\} = \boldsymbol{T}_d(a_i) = \operatorname{diag}\{\boldsymbol{\tau}_d(a_i)\}$, where $\boldsymbol{\tau}_d(a_i) = \{\tau_{d1}(a_i), \ldots, \tau_{dG}(a_i)\}$. We further assume that $\boldsymbol{T}_d(a)$ evolves smoothly along the covariate domain which allows the amount of variation attributed to each component $\varphi_{dg}(a, r, \omega)$ to vary smoothly as well. For covariance estimation, we target the smooth variance components through their corresponding precision matrix $\boldsymbol{T}_d^{-1}(a) = \boldsymbol{\Gamma}_d(a) = \operatorname{diag}\{\boldsymbol{\gamma}_d(a)\}$, where $\boldsymbol{\gamma}_d(a) = \{1/\tau_{d1}(a_i), \ldots, 1/\tau_{dG}(a_i)\} = \{\gamma_{d1}(a), \ldots, \gamma_{dG}(a)\}$. To estimate the smooth precision components, we project $\gamma_{dg}(a)$ onto a suitable basis, $\gamma_{dg}(a) = \sum_{p=1}^{P} \upsilon_{dgp} \psi_p(a)$, where υ_{dgp} are scalar precision components that act as basis weights and $\{\psi_p(a)_{p=1,\ldots,P}\}$ are basis functions spanning the covariate domain (e.g. B-splines). The dimension P of the basis functions is chosen to be sufficiently large to capture changes in the variance components along the covariate dimension. Given previous estimates for $\boldsymbol{\eta}_{di}(a)$ and $\boldsymbol{\varphi}_{dg}(a)$, estimates of the covariate-dependent variance components $\boldsymbol{\tau}_d(a)$ and measurement error variance σ_d^2 are obtained using REML methods [41].

The assumption that the variance components evolve smoothly over the covariate domain resolves several challenges that emerge when modeling the covariate-dependent variance components. First, the estimation procedure is able to borrow strength across the covariate-domain when modeling variation, a necessity when specific covariate values may only be observed once as in our motivating data. Second, we are able to project the precision components onto a low-rank basis of smooth functions which induces regularization and control over the speed at which $\boldsymbol{\tau}_d(a)$ is allowed to vary. Alternatively, a projection based approach would be less computationally burdensome with estimates of the subject-specific scores obtained directly by numerical integration, $\hat{\xi}_{dig}(a_i) = \langle Z_{di}(a_i, r, \omega), \hat{\varphi}_{dg}(a_i, r, \omega) \rangle$ and their corresponding variance components calculated empirically, $\hat{\tau}_{dg}(a) = \operatorname{cov}\{\hat{\xi}_{dig}(a)\}$, but the resulting estimates are unstable due to the limited number of observations at each point along the covariate domain. Therefore, despite the added compute time, our proposed linear mixed effects framework is better suited for providing covariate-adjustments to the region-referenced functional process in a controlled and principled manner.

The estimated covariate-dependent variance components are used to choose the number of eigencomponents included in the CA-HPCA decomposition where G'_d denotes a set of eigencomponents such that the total fraction of variance explained ($FVE_{dG'_d}$) is greater than 0.8 for each group

$d = 1, \ldots, D$. We recommend starting with a larger number $G_d = K_d L_d$ of covariate-dependent tensor components, $\{\varphi_{dg}(a_i, r, \omega) : 1, \ldots, G_d\}$, in the mixed effects modeling used for the estimation of the covariate-dependent variance components, $\{\tau_{dg}(a_i) : g = 1, \ldots, G_d\}$, and then reduce or add components as appropriate to determine the final value of G'_d. In order to estimate the group-specific total fraction variance explained via the G'_d covariate-dependent tensor components, we consider the quantity,

$$FVE_{dG'_d} = \frac{\int\{\sum_{i=1}^{n_d} \sum_{g=1}^{G'_d} \hat{\tau}_{dg}(a_i)\}da}{\int[\sum_{i=1}^{n_d}\{||\widehat{Y}_{di}^c(a_i, r, \omega)|| - R\int \hat{\sigma}_d^2 da\}]da},$$

where $||f(a_i, r, \omega)||^2 = \sum_{r=1}^{R} \int f(a_i, r, \omega)^2 d\omega$. Note that the above formulation utilizes covariate-dependent variance components estimates $\hat{\tau}_{dg}(a)$ and $\hat{\sigma}_d^2$ obtained from the proposed mixed effects model to calculate the ratio of the variance in the G'_d eigencomponents to the total variation in the observed data $Y_{di}(a_i, r, \omega)$ without measurement error. The denominator of $FVE_{dG'_d}$ does not use variation in a large number of tensor components to estimate the total variation in the observed data due to computational costs in fitting the proposed mixed effects model, but instead uses the two-dimensional norm of the demeaned data minus measurement error variance, similar to the approach by Chen et al. [7]. Consequently, when the measurement error variance is overestimated and scaled by $R\int \hat{\sigma}_d^2 da$, $FVE_{dG'_d}$ may exceed 1. Once G'_d is defined, the subject-specific scores can be obtained using their best linear unbiased predictor (BLUP),

$$
\begin{aligned}
\hat{\xi}_{dig}(a_i) &= E\{\xi_{dig}(a_i)|\boldsymbol{Y}_{di}\} \\
&= \hat{\tau}_{dg}(a_i)\hat{\boldsymbol{\varphi}}_{dg}(a_i)^T \widehat{\boldsymbol{\Sigma}}_{\boldsymbol{Y}_{di}}^{-1}\{\boldsymbol{Y}_{di}(a_i) - \hat{\boldsymbol{\eta}}_{di}(a_i)\},
\end{aligned}
$$

where $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{Y}_{di}} = \sum_{g=1}^{G'_d} \hat{\tau}_{dg}(a_i)\hat{\boldsymbol{\varphi}}_{dg}(a_i)\hat{\boldsymbol{\varphi}}_{dg}(a_i)^T + \hat{\sigma}_d^2 \boldsymbol{I}$. Predictions of subject-specific trajectories $\widehat{\boldsymbol{Y}}_{di}(a_i)$ may be formed as in (1) using estimated components. Asymptotic theory supporting consistent estimation of the variance components and subject-specific scores is discussed in [43]. For CA-HPCA, this estimation relies on the assumption of weak separability which ensures that the subject-specific scores are uncorrelated and thus the variance components form a diagonal matrix.

3.2 Inference via parametric bootstrap

Inference in the form of hypothesis testing and pointwise confidence intervals can be performed via a parametric bootstrap based resampling from the estimated CA-HPCA model components. To test the null hypothesis that all groups have equal means in the region r across the entire covariate domain, i.e. $H_0 : \eta_d(a, r, \omega) = \eta(a, r, \omega)$ for $d = 1, \ldots D$, we propose a parametric bootstrap procedure based on the CA-HPCA decomposition. The proposed parametric bootstrap generates outcomes based on the estimated model components under the null hypothesis for region r as $Y_{di}^b(a_i, r, \omega) =$

$\hat{\eta}(a_i, r, \omega) + \sum_{g=1}^{G'_d} \xi_{dig}^b(a_i)\hat{\varphi}_{dg}(a_i, r, \omega) + \epsilon_{di}^b(a_i, r, \omega)$ and as $Y_{di}^b(a_i, r', \omega) = \hat{\eta}_d(a_i, r', \omega) + \sum_{g=1}^{G'_d} \xi_{dig}^b\hat{\varphi}_{dg}(a_i, r', \omega) + \epsilon_{di}^b(a_i, r', \omega)$ in the other regions $r' \neq r$ not considered under the null, where subject-specific scores and measurement error are sampled from $\xi_{dig}^b(a_i) \sim \mathcal{N}\{0, \hat{\tau}_{dg}(a_i)\}$ and $\epsilon_{di}^b(a_i, r, \omega) \sim \mathcal{N}(0, \hat{\sigma}_d^2)$, respectively. The proposed test statistic $T_r = [\sum_{d=1}^D \int \int \{\hat{\eta}_d(a, r, \omega) - \hat{\eta}(a, r, \omega)\}^2 dad\omega]^{1/2}$ is based on the norm of the sum of square-integrated departures of the estimated group-region shifts $\hat{\eta}_d(a, r, \omega)$ from the estimate of the common shift across groups, $\hat{\eta}(a, r, \omega)$. The common region shift estimate $\hat{\eta}(a, r, \omega)$ under the null is set to the point-wise average of the group-region shift estimates, $\hat{\eta}_d(a, r, \omega)$, $d = 1, \ldots, D$. We utilize the proposed parametric bootstrap to estimate the distribution of the test statistic T_r which can be used to evaluate the null hypothesis across the covariate domain. Presented below is the algorithm for the proposed bootstrap test.

Algorithm 2 Bootstrap Test

For a fixed region, $r \in \{1, \ldots, R\}$, perform the following:

1. Generate B parametric bootstrap samples with sample size and age distribution in each group identical to the observed data.

2. For the bth parametric bootstrap sample, calculate the test statistic

$$T_r^b = \sqrt{\sum_{d=1}^D \int \int \{\hat{\eta}_d^b(a, r, \omega) - \hat{\eta}^b(a, r, \omega)\}^2 dad\omega},$$

where $\hat{\eta}_d^b(a, r, \omega)$ and $\hat{\eta}^b(a, r, \omega)$ are both estimated based on the bth bootstrap sample.

3. Use $(1/B)\sum_{b=1}^B \mathbb{I}(T_r^b > T_r)$ to estimate the p-value where $\mathbb{I}(\cdot)$ denotes the indicator function and T_r is the test statistic from the original sample.

The bootstrap test described above can be extended to test the null hypothesis that all groups have equal means in the region r for a fixed covariate value $a^* \in \mathcal{A}$, i.e. $H_0 : \eta_d(a^*, r, \omega) = \eta(a^*, r, \omega)$ for $d = 1, \ldots D$. This extension can be used to test for group differences at particular covariate values, for example at earlier or later developmental stages. Outcomes are generated as described above but the test statistic $T_r(a^*)$ is calculated at a fixed covariate value $a^* \in \mathcal{A}$ rather than integrated across the covariate domain. To generate point-wise confidence intervals for estimates of $\hat{\eta}_d(a, r, \omega)$, repeat the above parametric bootstrap procedure but instead generate outcomes from the model $Y_{di}^b(a_i, r, \omega) = \hat{\eta}_d(a_i, r, \omega) + \sum_{g=1}^{G'_d} \xi_{dig}^b\hat{\varphi}_{dg}(a_i, r, \omega) + \epsilon_{di}^b(a_i, r, \omega)$. At each iteration of the bootstrap, estimate $\hat{\eta}_d^b(a, r, \omega)$ from the simulated data and then form point-wise confidence intervals based on percentiles of the estimated bootstrap group-region mean functions across iterations as a function of a, r and ω, $\{\hat{\eta}_{dg}^b(a, r, \omega) : b = 1, \ldots, B\}$. The

two versions of the bootstrap test as well as point-wise confidence intervals will be utilized for data analysis in Section 4 and evaluated via simulations in Section 5.

4. APPLICATION TO THE 'EYES-OPEN' PARADIGM DATA

4.1 Data structure and methods

In our motivating data application, EEG signals were sampled at 500 Hz for two minutes from a 128-channel HydroCel Geodesic Sensor Net on 58 ASD and 39 TD children aged 25 to 146 months old (diagnostic groups were age matched). EEG recordings were collected during an 'eyes-open' paradigm in which bubbles were displayed on a screen in a sound-attenuated room to subjects at rest [14]. We describe the dataset in our previous work and present an abbreviated description here and direct the reader to Scheffler et al. [36] for technical details related to pre-processing and data acquisition. EEG data for each subject is interpolated down to a standard 10–20 system 25 electrode montage ($R = 25$, Figure 1(b)) using spherical interpolation as detailed in Perrin et al. [34], producing 25 electrodes with continuous EEG signal. Spectral density estimates for each electrode were obtained on the first 38 seconds of artifact free EEG data using Welch's method with two second Hanning windows and 50 percent overlap [40], where 38 seconds constitutes the minimum amount of artifact free data across subjects. Thus, for each subject the electrode-specific spectral estimates form an instance of region-referenced functional data. Given that our primary interest is to model the alpha spectrum as a form of functional data, we restrict our analysis to the alpha spectral band ($\Omega = (6 \text{ Hz}, 14 \text{ Hz})$) which due to the sampling scheme has a frequency resolution of .25 Hz resulting in $|\Omega| = 33$ functional grid points. The spectral density within this band is normalized to a unit area (through division of by its integral over Ω) to better facilitate comparisons across electrodes and subjects.

We employ the CA-HPCA decomposition to model the alpha spectrum which allows both the group-region mean functions and total variation to change across development. Estimation for the CA-HPCA procedure is carried out as described in Section 3.1. Smooths of the group-region mean functions $\eta_d(r, a, \omega)$ and covariate-dependent functional marginal covariances $\boldsymbol{\Sigma}_{d,\Omega}(a, \omega, \omega')$ are obtained using tensor bases formed from marginal penalized cubic B-splines (with 10 and 4 degrees of freedom in the functional and covariate domains, respectively) and second degree difference penalties along each dimension. Smooths of the precision components $\gamma_{dg}(a)$ for the linear mixed effects model are estimated by projection onto cubic B-splines with 4 degrees of freedom in the covariate domain. Penalty parameters and variance components for the group-region mean functions, covariate-dependent functional marginal covariances, and linear mixed effects model are selected via REML and models are fit using the `gam` and `bam` functions from
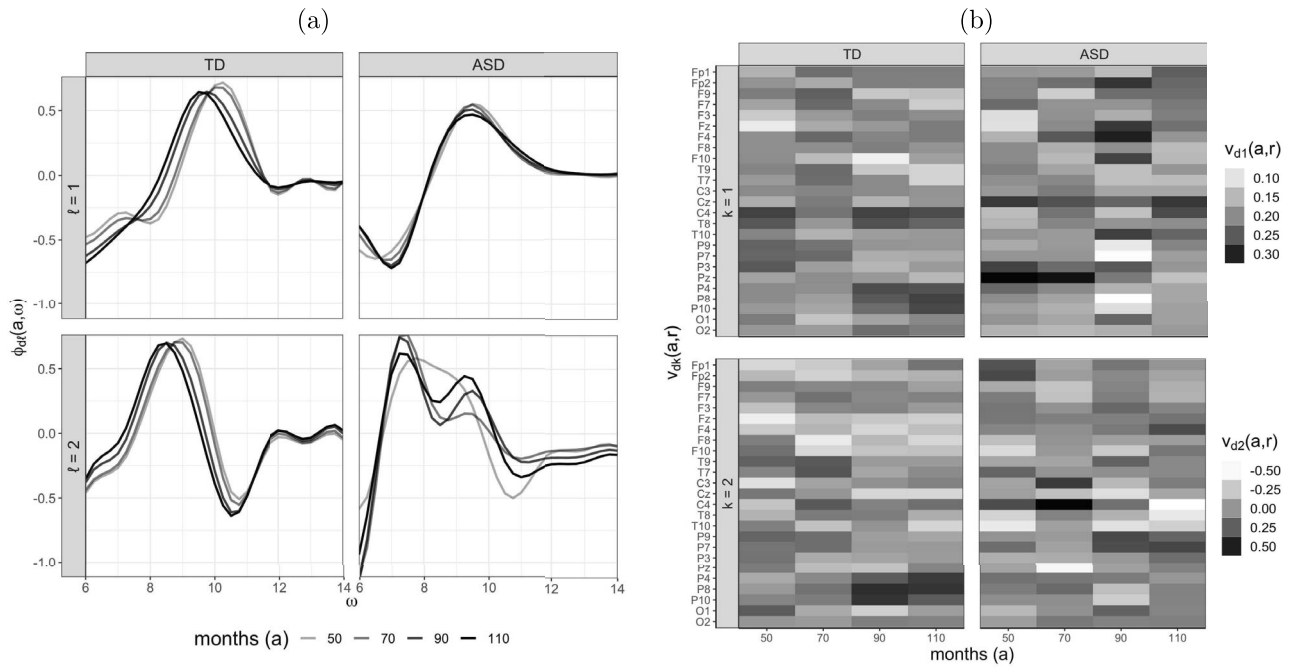
Figure 2. (a) Estimated first and second leading covariate-dependent eigenfunctions $\phi_{d1}(a,\omega)$ and $\phi_{d2}(a,\omega)$ at $a = 50, 70, 90, 110$ months (darker lines correspond to older children). (b) Estimated first and second leading covariate-dependent eigenvectors $v_{d1}(a,r)$ and $v_{d2}(a,r)$ at $a = 50, 70, 90, 110$ months. Shading corresponds to the weight of each element in the eigenvector.

mgcv (version 1.8-28). Smooths of the covariate-dependent regional marginal covariance $\{\boldsymbol{\Sigma}_{d,\mathcal{R}}(a)\}_{r,r'}$ are estimated as in Section 3.1 using a Gaussian kernel with Nadaraya–Watson estimates obtained using the ksmooth function from stats (version 3.6.1) and bandwidth selection performed via LOSOCV. The parametric bootstrap procedure utilized 200 bootstrap samples. All estimation was performed on a 2.8 GHz 6-Core Intel Core i7 processor operating the R software environment (version 3.6.1).

4.2 Data analysis results

We present the results from our application of the CA-HPCA decomposition to the EEG data. While the main focus of our analysis is to characterize differences in alpha spectral dynamics between TD and ASD children over the course of development via inference on the group-region mean functions, we begin by briefly discussing the covariate-dependent marginal eigencomponents produced by the decomposition. Figure S1 displays the marginal FVE of ordered eigencomponents for the covariate-dependent regional and functional marginal covariances that explain at least 90% of the marginal FVE across the covariate domain. The marginal FVE attributed to each component is relatively constant over development. The leading (five and six) covariate-dependent regional marginal eigenvectors and (four and four) covariate-dependent functional marginal eigenfunctions are collectively found to explain (1.006 and

0.895) of the total FVE ($FVE_{dG'_d}$) in the (TD and ASD) groups, respectively. Recall, the total FVE may exceed 1 due to slight overestimation of measurement error as described in Section 3.1.

In the functional dimension along the covariate domain, the leading covariate-dependent marginal eigenfunctions $\phi_{d1}(a,\omega)$ (Figure 2 (a), top row) display patterns of variation that capture individual differences between initial alpha power at 6 Hz and intermediate alpha power at 10 Hz. Note, the frequency location of maximal alpha variation decreases as age increases in the TD group but remains relatively constant across development in ASD group. The second leading covariate-dependent marginal eigenfunctions $\phi_{d2}(a,\omega)$ (Figure 2 (a), bottom row), identifies different sources of alpha variation in the TD and ASD groups. Referring back to the mean alpha spectral densities displayed in Figure 1 (a), both diagnostic groups display a dip before 10 Hz that is followed by a peak after 10 Hz, though this is much more pronounced in the TD group. For the TD group, the second eigenfunction can be interpreted as differences in alpha power below and above the prominent peak at 10 Hz, whereas in the ASD group the second eigenfunction captures difference between initial alpha power at 6 Hz and intermediate alpha power between 7 and 9 Hz. In the TD group, the location of variation in the second eigenfunction shifts horizontally from higher to lower frequencies as age increases but no covariate trend is detectable in the ASD group. The first two leading covariate-
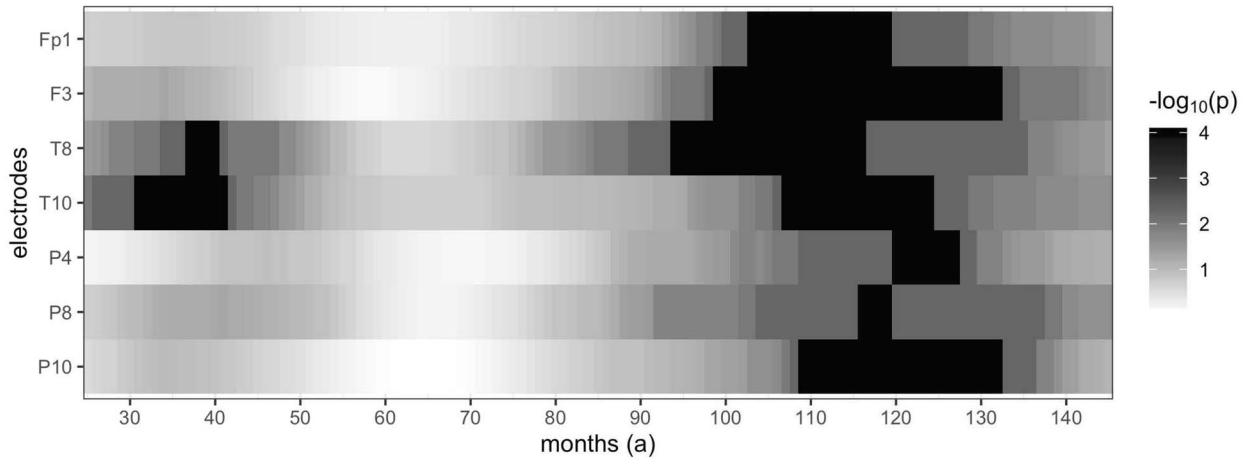
Figure 3. The results from the secondary analysis of seven electrodes identified as showing significant differences at some point across development. A parametric bootstrap test is conducted for each month between 25 and 145 months from the CA-HPCA model of the alpha spectrum. P-values are transformed to the $-\log_{10}$ scale to better stratify results where darker colors correspond to more significant differences ($-\log_{10}(p) > 1.30$ denote significance at level $\alpha = .05$).

dependent marginal eigenfunctions together explain at least 65% of the variation in the covariate-dependent functional marginal covariances in each diagnostic group.

In the regional dimension along the covariate domain, the first leading covariate-dependent marginal eigenvectors $v_{d1}(a, r)$ (Figure 2 (b), top row) display maximal variation in the (central and posterior regions) at younger ages with a shift to the (posterior and central regions) at older ages in the (TD and ASD groups), respectively. In the second leading covariate-dependent marginal eigenvector $v_{d2}(a, r)$ (Figure 2 (b), bottom row) the TD group shows maximal variation in the frontal region at younger ages with a shift to the the posterior at older ages while the ASD group exhibits maximal variation that alternates between the temporal and central regions over development. The first two covariate-dependent marginal eigenvectors together explain at least 70% of the variation in the covariate-dependent regional marginal covariances in each diagnostic group.

In the covariate dimension, the four leading covariate-dependent variance components $\tau_{dg}(a)$ (Figure S2) fluctuate as a function of age where the leading component corresponding to the covariate-dependent tensor component $\varphi_{d1}(a, r, \omega) = v_{d1}(a, r)\phi_{d1}(a, \omega)$ is responsible for the vast majority of variation in the ASD group but not the TD group. The variance components in the TD group vary along the covariate domain whereas the variance components in the ASD group are relatively constant apart from edge effects at the boundary of the covariate domain. In summary, the covariate-dependent eigencomponents and their associated variance components for the two diagnostic groups show differing patterns of variation that modulate distinctly along the covariate domain, confirming our need to model covariate-dependent heteroscedasticity in each group separately.

To test for differences between diagnostic groups in the alpha spectrum across development, we utilize the parametric bootstrap procedures described in Section 3.2. For each electrode r, we test the null hypothesis that the TD and ASD group-region mean functions are equal across the entire covariate domain from 25 to 145 months which takes the form $H_0 : \eta_d(r, \omega, a) = \eta(r, \omega, a)$, $d = 1, 2$. To address the issue of multiple testing across electrodes, we utilize the procedure of Benjamini and Yekutieli [3], a less conservative alternative to Bonferroni correction which transforms p-values into q-values to control the false discovery rate (FDR). We define q-values less than 0.05 to be statistically significant and find significant differences in the frontal (Fp1, q=0.036; F3, q<0.001), temporal (T8, q=0.025; T10, q=0.036), and posterior regions (P4, q=0.025; P8, q=0.025; P10, q<0.001) between the TD and ASD groups across development (results for each electrode can be found in Table S2). To investigate which stages of development are driving the observed differences in these seven electrodes (Fp1, F3, T8, T10, P4, P8, and P10), we perform a secondary analysis and test the null hypothesis that the TD and ASD group-region mean functions are equal for a fixed age $a^* = 25, 26, \ldots, 144, 145$ months within each electrode which takes the form $H_0 : \eta_d(r, \omega, a^*) = \eta(r, \omega, a^*)$, $d = 1, 2$. Because this is a secondary analysis, rather than formally adjusting for multiple comparisons we examine unadjusted p-values and avoid emphasis of findings with nominal significance, instead highlighting regions of development where electrodes consistently violate the null hypothesis. Figure 3 displays the results of the hypothesis tests for the seven electrodes with p-values transformed to the $-\log_{10}$ scale to better stratify results where values greater than $-\log_{10}(0.05) = 1.30$ denote significance at level $\alpha = .05$. All seven electrodes display strong differences between diagnostic groups in the
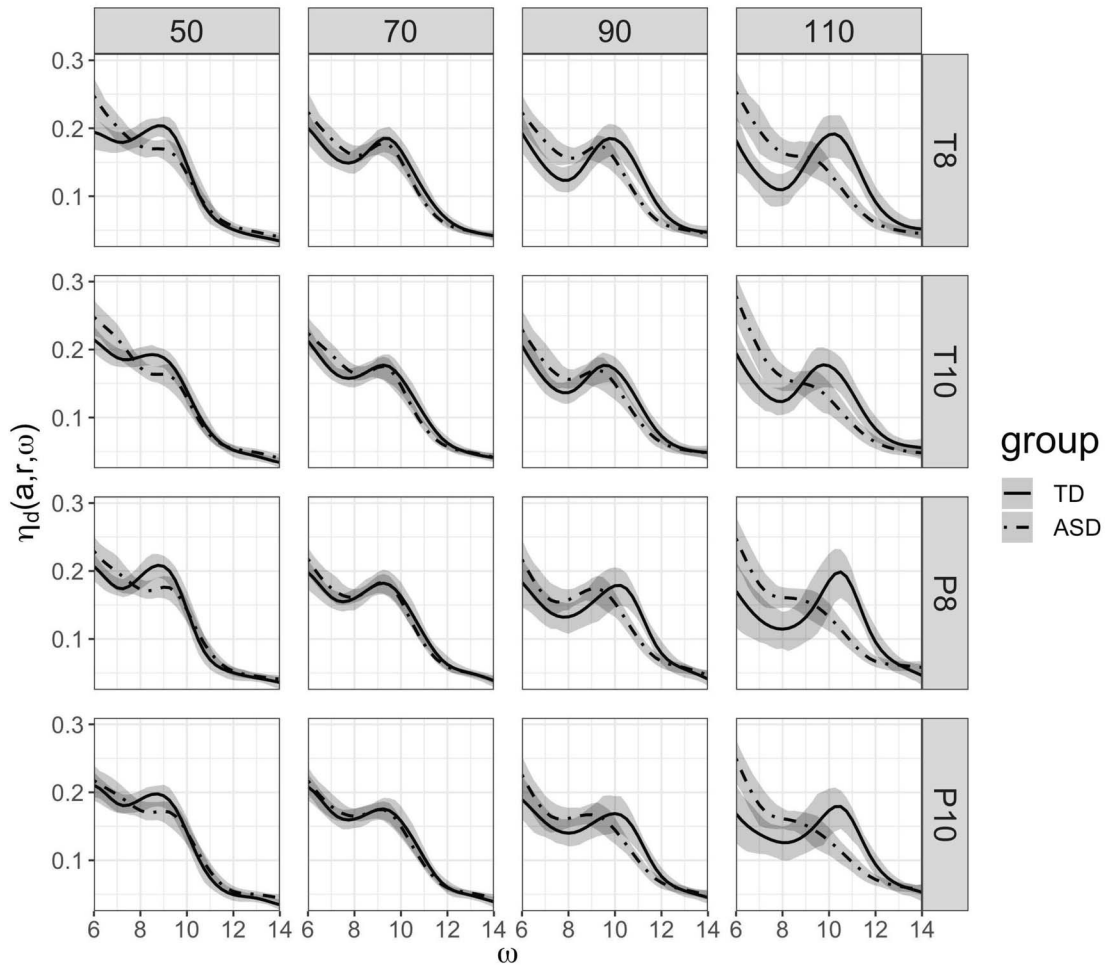
Figure 4. The estimated group-region mean functions $\eta_d(a, r, \omega)$ at ages $a = 50, 70, 90, 110$ months from the T8, T10, P8, and P10 electrodes from the CA-HPCA model of the alpha spectrum. Grey shading denotes 95% point-wise confidence intervals for estimates.

alpha spectrum between approximately 100 and 130 months with the two temporal electrodes showing differences at earlier ages between 30 and 50 months, as well.

Among the greatest visual differences in the group-region mean functions are observed in the T8, T10, P8, and P10 electrodes displayed in Figure 4 along with their 95% pointwise confidence intervals generated as described in Section 3.2. At all four electrodes, the TD group displays a well-defined peak in the alpha spectrum that shifts from 9 Hz to 11 Hz moving from 50 to 110 months, whereas the ASD group generally has less clearly-defined flat peaks that tend to center around 9 Hz throughout development. Differences in the estimated group-region mean functions mirror the results found from the secondary analysis which examined group-differences month by month. For the T8 electrode, the point-wise confidence intervals between diagnostic groups separate for younger and older ages at 50, 90 and 110 months, while all four electrodes display separation in the point-wide confidence intervals at 110 months.

To assess the sensitivity of our results to developmental information, we include a naive analysis examining group differences in the alpha spectrum using the hybrid principal components analysis decomposition of Scheffler et al. [35] which ignores covariate information. Full details of the naive analysis are included in Web Appendix B of the Supporting Information and are summarized here. The naive analysis find six regions that display differential alpha spectral dynamics between the two diagnostic groups over the course of development, four of which are not found among the seven electrodes identified by the CA-HPCA decomposition. Collectively, this suggests that omitting covariate-information reduces power in our motivating analysis and may lead to misleading results due to model misspecification. In addition, by omitting covariate information there is no way to quantify at what point in development these particular regions differ significantly. When aggregated, the observations and inferences obtained from the CA-HPCA model components provide evidence for differences in both

the mean structure and patterns of covariation between the two diagnostic groups that shift and change over development highlighting the need to provide covariate-adjustments in modeling the region-referenced EEG data across a broad age range.

5. SIMULATION

We studied the finite sample properties of the proposed CA-HPCA model as well as the associated bootstrap derived group-level inference via extensive simulations. We summarize the results of the simulation study here and defer details of data generation and simulation evaluation to Web Appendix C of the Supporting information. We conducted 500 Monte Carlo runs for two sample sizes ($n_d = 50$ and 100) and two signal-to-noise ratios (SNRs = 4 and 10) for a total of four settings. The lower sample size is similar to the group sample sizes in our observed EEG data. To assess the performance of the proposed estimation algorithm in targeting the functional and vector components of CA-HPCA, we utilize normalized mean squared errors (MSE) and relative squared errors (RSE) based on norms of deviations of the estimated quantities from target quantities. In addition, we report the total fraction of variance explained (FVE), coverage properties of 95% confidence intervals for the group-region mean functions, and power of the proposed bootstrap procedure for testing differences both across the covariate domain and at fixed locations of the covariate domain.

Figure 5 displays estimated model components based on 500 Monte Carlo runs from the CA-HPCA simulation setup with the most challenging simulation settings $n_d = 50$ and $c = 4$ (low SNR). The estimated group-region mean functions with the 10th, 50th, and 90th percentile RSE from $d = 1$ and $r = 5$ (Figure 5 (a)) closely match the true curves across the covariate domain. The estimated covariate-dependent regional and functional marginal eigencomponents (Figure 5 (b, c)) are displayed from runs with RSE values at the 10th, 50th, and 90th percentiles, overlaid by their true quantities. Even at small sample sizes and low SNR, CA-HPCA captures the periodicity, phase, and magnitude of the true components. Occasionally, estimates of model components at the edge of the covariate domain do not capture phase shifts likely due to relative sparsity observations in covariate domain when $n_d = 50$.

Table 1 displays median, 10th, and 90th percentile RSE and normalized MSE values based on 500 Monte Carlo runs corresponding to the estimated CA-HPCA components from all four simulation settings. Given that normalized measures of RSE and MSE were used, we report percentiles for model component(s) over all Monte Carlo runs combined across groups (and subjects in the case of subject-level predictions). More specifically, while performance measures for $v_{dk}(a, r)$, $\phi_{d\ell}(a, \omega)$, σ_d^2, and $FVE_{dG'_d}$ are reported over $D \times 500$ Monte Carlo runs, measures for $\eta_d(a, r, \omega)$, $Y_{di}(a_i, r, \omega)$, $\tau(a)_{d,k\ell}$, and coverage are reported
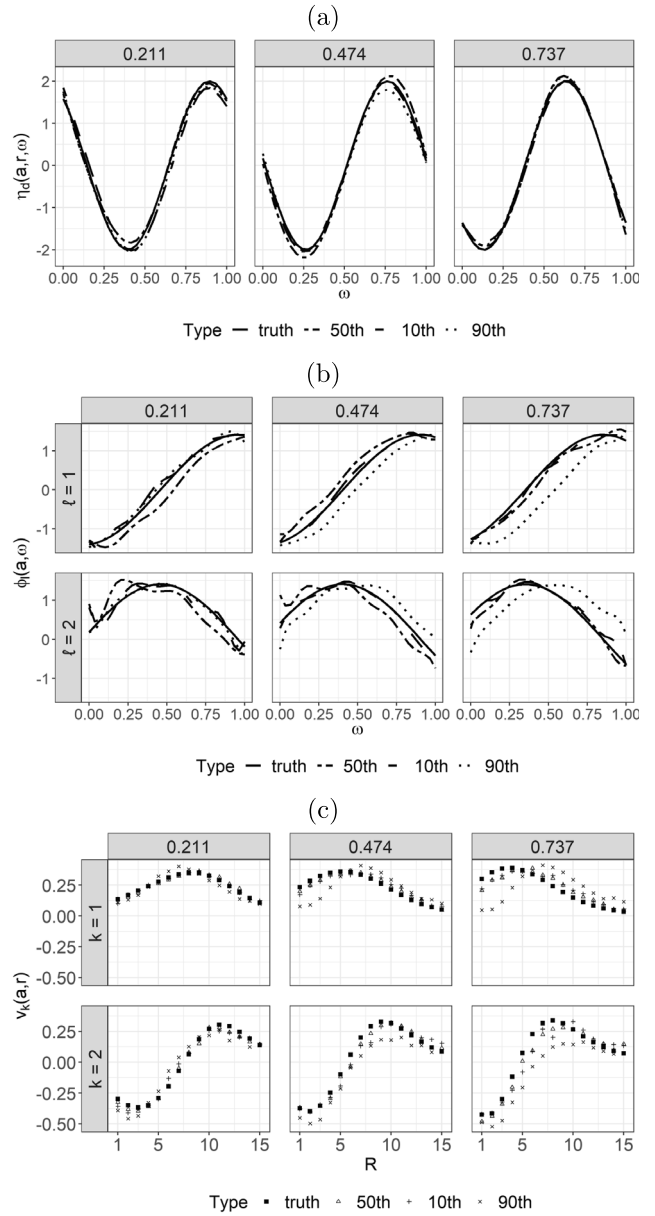

Figure 5. The true and estimated (a) group-region mean functions $\eta_d(a, r, \omega)$ for $d = 1$ and $r = 5$, (b) two leading covariate-dependent functional marginal eigenfunctions $\phi_{d1}(a, \omega)$ and $\phi_{d2}(a, \omega)$, (c) two leading covariate-dependent regional marginal eigenvectors $v_{d1}(a, r)$ and $v_{d2}(a, r)$ for $a = 0.211, 0.474, 0.737$ corresponding to the 10th, 50th, and 90th percentile relative squared error (RSE) values based on 500 Monte Carlo runs from the CA-HPCA simulation design at $n_d = 50$ and low signal-to-noise ratio (SNR).

over $D \times R \times 500$, $D \times n_d \times 500$, $D \times K_d \times L_d \times 500$, and $D \times R \times |\mathcal{A}| \times 500$ Monte Carlo runs, respectively.

Overall, the RSEs for all model components decrease with higher sample size and SNR. The predicted subject-

	Low SNR				High SNR			
	$n_d = 50$		$n_d = 100$		$n_d = 50$		$n_d = 100$	
$\eta_d(a,r,\omega)$	0.017	(0.007, 0.035)	0.009	(0.004, 0.018)	0.016	(0.006, 0.034)	0.008	(0.003, 0.017)
$Y_{di}(a_i,r,\omega)$	0.173	(0.140, 0.219)	0.173	(0.139, 0.220)	0.079	(0.062, 0.102)	0.078	(0.062, 0.102)
$v_{d1}(a,r)$	0.086	(0.024, 0.233)	0.045	(0.017, 0.130)	0.082	(0.022, 0.236)	0.042	(0.014, 0.097)
$v_{d2}(a,r)$	0.153	(0.069, 0.278)	0.074	(0.04, 0.159)	0.137	(0.061, 0.268)	0.067	(0.035, 0.128)
$\phi_{d1}(a,\omega)$	0.073	(0.026, 0.151)	0.048	(0.03, 0.097)	0.066	(0.031, 0.144)	0.048	(0.028, 0.088)
$\phi_{d2}(a,\omega)$	0.075	(0.031, 0.148)	0.052	(0.032, 0.098)	0.065	(0.033, 0.140)	0.049	(0.03, 0.091)
$\tau_{d,k\ell}(a)$	0.150	(0.031, 0.989)	0.061	(0.021, 0.975)	0.105	(0.032, 0.243)	0.053	(0.019, 0.141)
σ_d^2	0.056	(0.002, 0.320)	0.040	(0.002, 0.205)	0.067	(0.002, 0.430)	0.042	(0.001, 0.238)
$FVE_{d,k\ell}$	0.982	(0.962, 1.005)	0.992	(0.974, 1.006)	0.971	(0.947, 0.994)	0.985	(0.971, 1.001)
coverage	0.892	(0.758, 0.988)	0.952	(0.802, 0.998)	0.940	(0.823, 0.998)	0.958	(0.828, 0.998)

level curves $Y_{di}(a_i,r,\omega)$ are most sensitive to changes in SNR, as expected, while the RSEs for the eigencomponents, $v_{dk}(a,r)$, $\phi_{d\ell}(a,\omega)$ and $\tau_{d,k\ell}(a)$, are more sensitive to changes in sample size rather than SNR, suggesting that the estimation procedure effectively corrects for measurement error when obtaining the marginal covariances. The MSE for σ_d^2 was extremely small and did not follow a trend with respect to sample size or SNR. Across simulation designs, the total fraction of variance explained, FVE_{d,G'_d}, almost always approach 1.00 due to the compact number of marginal eigencomponents used to generate the data. Given that calculation of FVE_{d,G'_d} depends on estimates of the variance components and the two-dimensional norm of the demeaned observed data, the calculated values of FVE_{dG_d} may exceed 1.00 in some instances. For all simulation settings except the lowest sample size and SNR, the median coverage probabilities for the point-wise confidence intervals of the group-region mean functions approach their nominal level of 95%. For the hypothesis test defined across the covariate domain, the level of the parametric bootstrap test was approximately .05 for $n_d = 100$ and the power of the test generally increases faster with larger sample sizes (Table S1). For the hypothesis test at fixed locations of the covariate domain, the level of the parametric bootstrap test was slightly above .05 across the covariate domain, particularly at the smaller sample size $n_d = 50$. The power across fixed locations of the covariate domain also increased with sample size (Figure S4). Further discussion of the power analysis can be found in Web Appendix C of the Supporting Information.

6. DISCUSSION

We proposed a covariate-adjusted hybrid principal components analysis (CA-HPCA) which decomposes region-referenced functional data and accounts for covariate-dependent heteroscedasticity by assuming the high-dimensional covariance structure is weakly separable conditional on observed covariates. The proposed estimation procedure develops computationally efficient fast-covariance smoothers that incorporate covariate-dependence when estimating marginal covariances as well as a mixed effects framework which admits inference along the covariate-domain via parametric bootstrap sampling of estimated model components. As with any model, verifying key assumptions is necessary for principled inference, namely validating the assumption of weak separability conditional on observed covariates as well as joint normality of the subject-specific scores and measurement error variance in the linear mixed effects model. Application of CA-HPCA to region-referenced EEG data collected on TD and ASD children revealed that the alpha spectrum changes over development both in terms of mean structure and patterns of covariation. Further, inference based on the CA-HPCA decomposition revealed significant differences in alpha spectral dynamics between the two diagnostic groups, particularly at younger and older ages. The CA-HPCA decomposition was developed to model EEG data over a broad developmental range, the procedure may be applied to other settings where high-dimensional data is expected to exhibit differential covariation as a function of observed covariates.

REFERENCES

[1] BACKENROTH, D., GOLDSMITH, J., HARRAN, M. D., CORTES, J. C., KRAKAUER, J. W. and KITAGO, T. (2018). Modeling Motor Learning Using Heteroscedastic Functional Principal Components Analysis. *Journal of the American Statistical Association* **113** 1003–1015. MR3862335

[2] BALADANDAYUTHAPANI, V., MALLICK, B. K., HONG, M. Y., LUPTON, J. R., TURNER, N. D. and CARROLL, R. J. (2007). Bayesian Hierarchical Spatially Correlated Functional Data Analysis with Application to Colon Carcinogenesis. *Biometrics* **64** 64–73. MR2422820

[3] BENJAMINI, Y. and YEKUTIELI, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Ann. Statist.* **29** 1165–1188. MR1869245

[4] BRUCE, S. A., HALL, M. H., BUYSSE, D. J. and KRAFTY, R. T. (2018). Conditional adaptive Bayesian spectral analysis of nonstationary biomedical time series. *Biometrics* **74** 260–269. MR3777946

[5] CARDOT, H. (2007). Conditional Functional Principal Components Analysis. *Scandinavian Journal of Statistics* **34** 317–335. MR2346642

[6] CEDERBAUM, J., SCHEIPL, F. and GREVEN, S. (2018). Fast symmetric additive covariance smoothing. *Computational Statistics & Data Analysis* **120** 25–41. MR3742206

[7] CHEN, K., DELICADO, P. and MÜLLER, H. G. (2016). Modelling function-valued stochastic processes, with applications to fertility dynamics. *Journal of the Royal Statistical Society. Series B (Methodological)* **79** 177–196. MR3597969

[8] CHEN, K. and MÜLLER, H. G. (2012). Modeling Repeated Functional Observations. *Journal of the American Statistical Association* **107** 1599–1609. MR3036419

[9] CHIOU, J.-M., CHEN, Y.-T. and YANG, Y.-F. (2014). Multivariate functional principal components analysis: a normalization approach. *Statistica Sinica* **24** 1571–1596. MR3308652

[10] CHIOU, J.-M., MÜLLER, H.-G. and WANG, J.-L. (2003). Functional Quasi-Likelihood Regression Models with Smooth Random Effects. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* **65** 405–423. MR1983755

[11] CORCORAN, A. W., ALDAY, P. M., SCHLESEWSKY, M. and BORNKESSEL-SCHLESEWSKY, I. (2018). Toward a reliable, automated method of individual alpha frequency (IAF) quantification. *Psychophysiology* **55** e13064.

[12] CRAINICEANU, C., STAICU, A. M. and DI, C. Z. (2009). Generalized Multilevel Functional Regression. *Journal of the American Statistical Association* **104** 1550–1561. MR2750578

[13] DI, C. Z., CRAINICEANU, C. M., CAFFO, B. S. and PUNJABI, N. M. (2009). Multilevel functional principal component analysis. *The Annals of Applied Statistics* **3** 458–488. MR2668715

[14] DICKINSON, A., DISTEFANO, C., SENTURK, D. and JESTE, S. S. (2018). Peak alpha frequency is a neural marker of cognitive function across the autism spectrum. *European Journal of Neuroscience* **47** 643–651.

[15] EDGAR, J. C., DIPIERO, M., MCBRIDE, E., GREEN, H. L., BERMAN, J., KU, M., LIU, S., BLASKEY, L., KUSCHNER, E., AIREY, M., ROSS, J. L., BLOY, L., KIM, M., KOPPERS, S., GAETZ, W., SCHULTZ, R. T. and ROBERTS, T. P. L. (2019). Abnormal maturation of the resting-state peak alpha frequency in children with autism spectrum disorder. *Human Brain Mapping* **40** 3288–3298.

[16] EDGAR, J. C., HEIKEN, K., CHEN, Y.-H., HERRINGTON, J. D., CHOW, V., LIU, S., BLOY, L., HUANG, M., PANDEY, J., CANNON, K. M., QASMIEH, S., LEVY, S. E., SCHULTZ, R. T. and ROBERTS, T. P. L. (2015). Resting-State Alpha in Autism Spectrum Disorder and Alpha Associations with Thalamic Volume. *Journal of Autism and Developmental Disorders* **45** 795–804.

[17] FIECAS, M. and OMBAO, H. (2016). Modeling the Evolution of Dynamic Brain Processes During an Associative Learning Experiment. *Journal of the American Statistical Association* **111** 1440–1453. MR3601700

[18] GIRALDO, R., DELICADO, P. and MATEU, J. (2010). Ordinary kriging for function-valued spatial data. *Environmental and Ecological Statistics* **18** 411–426. MR2832903

[19] GREVEN, S., CRAINICEANU, C. M., CAFFO, B. S. and REICH, D. S. (2010). Longitudinal functional principal component analysis. *Electronic Journal of Statistics* **4** 1022–1054. MR2727452

[20] HAPP, C. and GREVEN, S. (2018). Multivariate Functional Principal Component Analysis for Data Observed on Different (Dimensional) Domains. *Journal of the American Statistical Association* **113** 649–659. MR3832216

[21] HASENSTAB, K., SCHEFFLER, A., TELESCA, D., SUGAR, C. A., JESTE, S., DISTEFANO, C. and ŞENTÜRK, D. (2017). A multidimensional functional principal components analysis of EEG data. *Biometrics.* MR3713133

[22] JAQUES, J. and PREDA, C. (2014). Model-based Clustering for Multivariate Functional Data. *Computational Statistics and Data Analysis* **71** 92–106. MR3131956

[23] JIANG, C.-R. and WANG, J.-L. (2010). Covariate adjusted functional principal components analysis for longitudinal data. *Ann. Statist.* **38** 1194–1226. MR2604710

[24] KARHUNEN, K. (1946). Zur Spektraltheorie stochastischer Prozesse. *Ann. Acad. Sci. Fennicae, AI* **37** 1–37. MR0023012

[25] KRAFTY, R. T., ROSEN, O., STOFFER, D. S., BUYSSE, D. J. and HALL, M. H. (2017). Conditional Spectral Analysis of Replicated Multiple Time Series With Application to Nocturnal Physiology. *Journal of the American Statistical Association* **112** 1405–1416. MR3750864

[26] KUNDU, M. G., HAREZLAK, J. and RANDOLPH, T. W. (2016). Longitudinal functional models with structured penalties. *Statistical Modelling* **16** 114–139. MR3478713

[27] LIU, C., RAY, S. and HOOKER, G. (2016). Functional principal component analysis of spatially correlated data. *Statistics and Computing* 1–16. MR3687330

[28] LOEVE, M. (1946). Fonctions aléatoires à décomposition orthogonale exponentielle. *La Revue Scientifique* **84** 159–162. MR0017892

[29] LYNCH, B. and CHEN, K. (2018). A test of weak separability for multi-way functional data, with application to brain connectivity studies. *Biometrika* **105** 815–831. MR3877867

[30] MISKOVIC, V., MA, X., CHOU, C. A., FAN, M., OWENS, M., SAYAMA, H. and GIBB, B. E. (2015). Developmental changes in spontaneous electrocortical activity and network organization from early to late childhood. *NeuroImage* **118** 237–247.

[31] MORRIS, J. S., VANNUCCI, M., BROWN, P. J. and CARROLL, R. J. (2003). Wavelet-Based Nonparametric Modeling of Hierarchical Functions in Colon Carcinogenesis. *Journal of the American Statistical Association* **98** 573–583. MR2011673

[32] MORRIS, J. S. and CARROLL, R. J. (2006). Wavelet-based functional mixed models. *Journal of the Royal Statistical Society. Series B (Methodological)* **68** 179–199. MR2188981

[33] PARK, S. Y. and STAICU, A. M. (2015). Longitudinal functional data analysis. *Stat* **4** 212–226. MR3405402

[34] PERRIN, F., PERNIER, J., BERTRAND, O. and ECHALLIER, J. F. (1989). Spherical splines for scalp potential and current density mapping. *Electroencephalography and Clinical Neurophysiology* **72** 184–187.

[35] SCHEFFLER, A., TELESCA, D., LI, Q., ŞENTÜRK, D., SUGAR, C. A., DISTEFANO, C. and JESTE, S. (2018). Hybrid principal components analysis for region-referenced longitudinal functional EEG data. *Biostatistics.* MR4043850

[36] SCHEFFLER, A. W., TELESCA, D., SUGAR, C. A., JESTE, S., DICKINSON, A., DISTEFANO, C. and ŞENTÜRK, D. (2019). Covariate-adjusted region-referenced generalized functional linear model for EEG data. *Statistics in Medicine* **38** 5587–5602. MR4039196

[37] STAICU, A. M., CRAINICEANU, C. M. and CARROLL, R. J. (2010). Fast methods for spatially correlated multilevel functional data. *Biostatistics* **11** 177–194.

[38] VALDAS-HERNANDEZ, P. A., OJEDA-GONZALEZ, A., MARTINEZ-MONTES, E., LAGE-CASTELLANOS, A., VIRUES-ALBA, T., VALDAS-URRUTIA, L. and VALDES-SOSA, P. A. (2010). White matter architecture rather than cortical surface area correlates with the EEG alpha rhythm. *NeuroImage* **49** 2328–2339.

[39] WANG, J.-L., CHIOU, J.-M. and MÜLLER, H. G. (2016). Functional Data Analysis. *Annual Review of Statistics and Its Application* **3** 257–295.

[40] WELCH, P. D. (1967). The use of Fast Fourier Transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. *IEEE Transactions on audio and electroacoustics* **15** 70–73.

[41] WOOD, S. (2017). *Generalized additive models: an introduction with R.* Chapman and Hall/CRC. MR2206355

[42] XIAO, L., LI, C., CHECKLEY, W. and CRAINICEANU, C. (2018). Fast covariance estimation for sparse functional data. *Stat. Comput.* **28** 511–522. MR3761337

[43] YAO, F., MÜLLER, H. G. and WANG, J. L. (2005). Functional Data Analysis for Sparse Longitudinal Data. *Journal of the American Statistical Association* **100** 577–590. MR2160561

[44] ZHOU, L., HUANG, J. Z., MARTINEZ, J. G., MAITY, A., BALADANDAYUTHAPANI, V. and CARROLL, R. J. (2010). Reduced Rank Mixed Effects Models for Spatially Correlated Hierarchical Functional Data. *Journal of the American Statistical Association* **105** 390–400. MR2757206

[45] ZIPUNNIKOV, V., GREVEN, S., SHOU, H., CAFFO, B. S., REICH, D. S. and CRAINICEANU, C. M. (2014). Longitudinal high-dimensional principal components analysis with application to diffusion tensor imaging of multiple sclerosis. *The Annals of Applied Statistics* **8** 2175–2202. MR3292493

Aaron Wolfe Scheffler
Department of Epidemiology & Biostatistics
University of California
San Francisco
USA
E-mail address: aaron.scheffler@ucsf.edu

Abigail Dickinson
Department of Psychiatry and Biobehavioral Sciences
University of California
Los Angeles
USA
E-mail address: ADickinson@mednet.ucla.edu

Charlotte DiStefano
Department of Psychiatry and Biobehavioral Sciences
University of California
Los Angeles
USA
E-mail address: CDiStefano@mednet.ucla.edu

Shafali Jeste
Departments of Neurology and Pediatrics
Keck School of Medicine
University of Southern California
Los Angeles
USA
E-mail address: sjeste@chla.usc.edu

Damla Şentürk
Department of Biostatistics
University of California
Los Angeles
USA
E-mail address: dsenturk@ucla.edu