Robust conditional spectral analysis of replicated time series

ZEDA LI

Classical second-order spectral analysis, which is based on the Fourier transform of the autocovariance functions, focuses on summarizing the oscillatory behaviors of a time series. However, this type of analysis is subject to two major limitations: first, being covariance-based, it cannot captures oscillatory information beyond the second moment, such as time-irreversibility and kurtosis, and cannot accommodate heavy-tail dependence and infinite variance; second, focusing on a single time series, it is unable to quantify the association between multiple time series and other covariates of interests. In this article, we propose a novel nonparametric approach to the spectral analysis of multiple time series and the associated covariates. The procedure is based on the copula spectral density kernel, which inherits the robustness properties of quantile regression and does not require any distributional assumptions such as the existence of finite moments. Copula spectral density kernels of different pairs are modeled jointly as a matrix to allow flexible smoothing. Through a tensor-product spline model of Cholesky components of the conditional copula spectral density matrix, the approach provides flexible nonparametric estimates of the copula spectral density matrix as nonparametric functions of frequency and covariate while preserving geometric constraints. Empirical performance is evaluated in simulation studies and illustrated through an analysis of stride interval time series.

1. INTRODUCTION

Understanding frequency patterns of a time series is essential to addressing scientific questions in a variety of fields, including biomedical science, environmental science, geology, and engineering. With advances in technology, an increasing number of studies collect time series data across different subjects, which we refer to as replicated time series, to explore the connections between frequency patterns of the underlying processes and covariates of interest. Our motivating example comes from a gait maturation study, where the stride interval or the gait cycle duration of 50 healthy children are recorded. The primary goal of this study is to understand whether the dynamics of walking depends on age. Three examples of stride interval time series in this study are presented in Figure 1.

Classical spectral analysis and frequency domain methods rely on the spectral density, defined as the Fourier transform of the autocovariance functions, providing powerful tools to quantify oscillatory patterns in a single univariate or multivariate time series [36, 5, 4]. However, the classical spectral analysis approaches have two major limitations. First, being entirely based on the autocovariance functions, it hinders the scope of processes that can be accurately analyzed. For example, it is not robust against outliers and heavy-tails; it cannot capture important dynamic features such as skewness, kurtosis, and time-irreversibility. Second, focusing on a single time series instead of replicated time series, it provides no information about the association between spectra and covariates.

Recently, efforts have been made to address these two limitations. On one hand, quantile-related spectral quantities have been developed, including the Laplace periodogram [22, 30, 18], the quantile periodogram [31], the copula spectral density kernel [12, 26, 44], and the quantile spectral analysis for locally stationary time series [3]. However, they are designed for analysis of a single time series and are unable to quantify the associations between the replicated time series and covariates. On the other hand, methods for spectral analysis of replicated time series have also recently been proposed, which include semi-parametric functional mixed-effects models [28, 8], nonparametric methods [14], and adaptive Bayesian approaches [6, 29, 7, 33]. However, these methods have roots in the classical autocovariance-based spectral density, and thus are subject to the known limitations of the classical second-order spectral analysis.

To address the aforementioned two limitations, we introduce a novel nonparametric method to the spectral analysis of replicated time series and the associated covariates. Using the copula spectral density kernel introduced by [26], the proposed method is robust against outliers and heavy-tails and allows for inferences on skewness, kurtosis, and time-irreversibility. Through a tensor-product spline model of Cholesky components of the conditional copula spectral density matrix, the approach provides flexible nonparametric estimates of copula spectral density matrix as functions of frequency and covariate while allowing different levels of smoothness. Formulated in a fully Bayesian framework,
is defined as a single time series. In this section, we present a Bayesian smoothing spline model for time series from multiple subjects, which we discuss in Section 3. Simulation studies are provided in Section 4 to illustrate the advantages of the proposed method. The use of the procedure in real-world applications is illustrated in Section 5, where it is used to analyze stride interval time series from a gait maturation study. Conclusions and future directions of this work are covered in Section 6.

2. SINGLE TIME SERIES

The proposed approach for the copula spectral analysis of replicated time series from multiple subjects, which we discuss in Section 3, is based on Bayesian splines. To motivate the development of the estimation procedure, in this section, we present a Bayesian smoothing spline model for a single time series.

2.1 The copula spectral density kernel

Consider a strictly stationary process, \((X_t)_{t \in \mathbb{Z}}\), of which we observe a finite stretch \(X_1, \ldots, X_T\). Let \(F\) be the marginal distribution of \(X_t\), and \(q_\tau := \inf\{x \in \mathbb{R} : \tau \leq F(x)\}, \tau \in [0, 1]\), be the corresponding quantile function. If \(\tau \in [0, 1]\), then \(-\infty\) and \(\infty\) are possible values for \(q_\tau\) and we let \(\inf \emptyset = \infty\). The lag-\(k\) copula cross-covariance kernel is defined as

\[
\gamma^U_k(\tau_1, \tau_2) = \text{Cov}(I\{U_t \leq \tau_1\}, I\{U_{t-k} \leq \tau_2\}),
\]

where \(k \in \mathbb{Z}, \tau_1, \tau_2 \in [0, 1]\), \(U_t := F(X_t)\), and \(I\{\}\) is the indicator function. If the cross-covariance kernels \(\gamma^U_k(\tau_1, \tau_2)\) are absolutely summable, we can define the copula spectral density kernel as, for \(\omega \in \mathbb{R}, (\tau_1, \tau_2) \in [0, 1]^2\),

\[
f_{q_{\tau_1}, q_{\tau_2}}(\omega) = \sum_{k \in \mathbb{Z}} \gamma^U_k(\tau_1, \tau_2) \exp(-2\pi i \omega k).
\]

Note that \(f_{q_{\tau_1}, q_{\tau_2}}(\omega)\) is continuous and satisfies \(f_{q_{\tau_1}, q_{\tau_2}}(-\omega) = f_{q_{\tau_2}, q_{\tau_1}}(\omega)\), where \(\overline{f}\) is the conjugate of \(f\). The copula spectral density kernels were introduced by [12], which is a generalization to the quantile spectral densities of [18]. It can be shown that \(\gamma^U_k(\tau_1, \tau_2)\) is the probability for \(X_t\) to switch from the upper \(\tau_k\) tail to the lower \(\tau_k\) tail in \(k\) steps minus the corresponding probability for white noise. Correspondingly, the copula spectral density kernels provide information on the quantile-crossing, or tail-switching probabilities. Contrary to the classical second-order spectral density, which is based on the autocovariance, the copula spectral density kernel allows for a complete description of the pairwise copulas of a time series by accounting for much more than the autocovariance structure of a time series.

We estimate the copula spectral density kernels for a set of \(\tau = \{\tau_1, \ldots, \tau_p\} \in [0, 1]\) jointly. To accomplish this, we consider the copula spectral density matrix such that

\[
f(\omega) = \{f_{q_{\tau_j}, q_{\tau_k}}(\omega)\}_{j,k=1}^p.
\]

The copula spectral density matrix is a \(p \times p\) positive definite, periodic and Hermitian function of frequency such that \(f(\omega) = f(\omega + 1)\), and \(f(\omega) = f^*(\omega)\), where \(f^*(\omega)\) is conjugate transpose of \(f(\omega)\). Equivalently, real components are even functions \(\Re\{f(\omega)\} = \Re\{f(-\omega)\}\) and imaginary components are odd functions \(\Im\{f(\omega)\} = -\Im\{f(-\omega)\}\). A non-vanishing imaginary part for \(f_{q_{\tau_1}, q_{\tau_2}}(\omega)\) indicates that \(\gamma^U_k(\tau_1, \tau_2)\) differs from \(\gamma^L_k(\tau_1, \tau_2)\), and further implies that \(X_t\) is not time-reversible. This formulation, which allows for simultaneous estimation of pairwise copula spectral density kernels with a set of \(\tau\), has also been considered in [44].
2.2 Estimation

2.2.1 Likelihood of the clipped time series

Our nonparametric estimation of \( f(\omega) \) from an epoch of length \( T, X_1, \ldots, X_T \), begins by considering the discrete Fourier transform (DFT) of the so-called clipped time series [23, 18, 26]

\[
Y_{\tau,m} = T^{-1/2} \sum_{t=1}^{T} I\{\hat{F}_T(X_t) \leq \tau\} \exp(-2\pi i \omega_m t)
\]

\[
(2) \quad Y_{\tau,m} = T^{-1/2} \sum_{t=1}^{T} I\{R_{T:t} \leq T\tau\} \exp(-2\pi i \omega_m t),
\]

where \( \hat{F}_T(X_t) = T^{-1} \sum_{t=1}^{T} I\{X_t \leq x\} \) is the empirical marginal distribution function, \( R_{T:t} \) is the rank of \( X_t \) among \( X_1, \ldots, X_T \), \( \omega_m = m/T, m = 1, \ldots, M \) are the Fourier frequencies, and \( M = \lfloor (T - 1)/2 \rfloor \), which is the greatest integer that is less than or equal to \( (T - 1)/2 \). Let \( Y_m = (Y_{\tau,m}, \ldots, Y_{\tau,m})' \). Then, \( Y_m \) are approximately independent \( p \times p \)-dimensional mean-zero complex Gaussian random variables with variances \( f(\omega_m) \) for \( m = 1, \ldots, M \) when \( T \) is large [26], which leads to the following log-likelihood function

\[
\mathcal{L}(Y \mid f) \approx -\sum_{m=1}^{M} \{ \log [f(\omega_m)] + Y_m^* f^{-1}(\omega_m) Y_m \},
\]

where \( Y \) represents the collection of all discrete Fourier transforms of the clipped time series for a set of \( \tau \). Similar to the periodograms that correspond to the classical spectral density, the rank-based copula periodogram (CR-periodogram), \( Y_m Y_m^* \), provides approximately unbiased but noisy estimates of \( f(\omega_m) \). Consistent estimates can be obtained by smoothing the periodogram across frequency through weighted average [12, 26]. We, on the other hand, propose using Bayesian smoothing spline models to estimate \( f(\omega) \). Two aspects of the log-likelihood function should be noted. First, other types of periodograms, such as those in [30, 18, 12], can also be used. We choose to work on the CR-periodogram derived from the DFT of the clipped time series for its easy computation and theoretical justifications.

Second, the likelihood function is different from, although in a similar form, the Whittle likelihood [42] used in the traditional second-order autocovariance-based spectral analysis. The Whittle likelihood is particularly useful in the analysis of Gaussian time series since the distribution of a Gaussian process is completely characterized by its spectral density, but it is invalid if the sample size is small and the time series is non-Gaussian [10]. Although [25] proposed a correction to the Whittle likelihood for non-Gaussian time series, it is still based on the second-order spectral density and is subject to the known limitations. The likelihood function considered in (3), however, does not require any distributional assumptions such as Gaussianity.

2.2.2 Bayesian smoothing spline

To preserve the positive-definiteness while allowing for different levels of smoothness, we model \( f(\omega) \) through the modified Cholesky decomposition [11, 37] such that

\[
f^{-1}(\omega) = \Theta(\omega) \Psi^{-1}(\omega) \Theta(\omega)^*.
\]

where \( \Theta(\omega) = [\Theta_{jk}(\omega)]_{j,k=1,\ldots,p} \) is a unique \( p \times p \) lower triangular complex matrix with ones on the diagonal, and \( \Psi(\omega) = \text{diag}[\Psi_{11}(\omega), \ldots, \Psi_{pp}(\omega)] \) is a unique \( p \times p \) positive diagonal matrix. Then, there exist \( p^2 \) Cholesky components to estimate: \( \Re\{\Theta_{jk}(\omega)\} \) for \( j > k = 1, \ldots, p, \Im\{\Theta_{jk}(\omega)\} \) for \( j > k = 1, \ldots, p, \) and \( \Im jj(\omega) \) for \( j = 1, \ldots, p \). Since the diagonal terms of \( \Psi^{-1}(\omega) \) is positive, we model it in log scale.

Smoothing spline estimation balances the fit of a function to the observed data with a roughness-based measure of regularity. In the case of Gaussian observations, the Bayesian formulation of smoothing splines was first introduced by [24, 39]. For spectral analysis, we measure the roughness of \( f \) through the summation of integrated squared second derivatives of the Cholesky components, \( \mathcal{P}(\Re\{\Theta_{jk}(\omega)\}) = \int_{-1/2}^{1/2} \Re\{\Theta''_{jk}(\omega)\}^2 d\omega, \mathcal{P}(\Im\{\Theta_{jk}(\omega)\}) = \int_{-1/2}^{1/2} \Im\{\Theta''_{jk}(\omega)\}^2 d\omega, \) and \( \mathcal{P}(\log \Psi_{jj}) = \int_{-1/2}^{1/2} \log \Psi_{jj}''(\omega)^2 d\omega \). The real components are periodic even functions and we specify the prior distribution by decomposing \( \mathcal{P}(\Theta_{jk}(\omega)) \) into a linear part and a nonlinear part. The linear part of a periodic even function is a constant. For the nonlinear part, we consider the reproducing kernel of the seminorm defined by \( \mathcal{P} \) over smooth periodic even functions

\[
J(\omega_s, \omega_m) = -1/2 \left[ \kappa_4(\omega_s - \omega_m) + \kappa_4(\omega_s + \omega_m - [\omega_s + \omega_m]) \right],
\]

where \( \kappa_1(\omega) = \omega - 0.5, \kappa_2(\omega) = |\kappa_2^2(\omega) - 1/12|/2, \) and \( \kappa_4(\omega) = \lfloor \kappa_4^2(\omega)/2 + 7/240 \rfloor/24 \) are the scaled Bernoulli polynomials [40]. The imaginary components are periodic odd functions. Consequently, \( \Im\{\Theta_{jk}(0)\} = 0 \), and the linear part of imaginary components is null. The reproducing kernel of the seminorm defined by \( \mathcal{P} \) over smooth periodic odd functions is given by

\[
K(\omega_s, \omega_m) = -1/2 \left[ \kappa_4(\omega_s + \omega_m - [\omega_s + \omega_m]) - \kappa_4(\omega_s - \omega_m) \right].
\]

The use of the kernel \( J \) and \( K \) accounts for the geometric restrictions to improve performance at the boundaries while avoiding redundant information [27]. Our formulation is closely related to the approach of [29]. They considered a slightly different reproducing kernel in \( C^{(2)}(0, 1) \) derived from the standard Taylor expansion such that \( J(\omega_s, \omega_m) = \int_0^1 (\omega_s - \mu)_+ (\omega_m - \mu)_+ d\mu, \) where \( (\mu)_+ = \max(\mu, 0) \). The two formulations yield similar results for univariate smoothing.

Robust conditional spectral analysis of replicated time series
but our approach can be more naturally extended to constructing reproducing kernels on $[0, 1]^2$ [17, Section 2.4].

For the Bayesian smoothing spline model, the prior distributions for the Cholesky components are

$$
\Re\{\Theta_{jk}(\omega)\} = a_{rjk} + a_{rjk}\omega + \sum_{m=1}^{M} J(\omega, \omega_m)z_{rjk,m},
$$

$$
\Im\{\Theta_{jk}(\omega)\} = \sum_{m=1}^{M} K(\omega, \omega_m)z_{rjk,m},
$$

$$
\log\{\Psi_{jj}^{-1}(\omega)\} = a_{djj} + a_{djj}\omega + \sum_{m=1}^{M} J(\omega, \omega_m)z_{djj,m},
$$

where we assume $a_{rjk} = (a_{rjk1}, a_{rjk2})' \sim N(0, \sigma_1 I_2)$, $a_{djj} = (a_{djj1}, a_{djj2})' \sim N(0, \sigma_2 I_2)$, $z_{rjk} = (z_{rjk1}, \ldots, z_{rjkM})' \sim N(0, \lambda_{rjk}^2 I_M^{-1})$, $z_{djj} = (z_{djj1}, \ldots, z_{djjM})' \sim N(0, \lambda_{djj}^2 I_M^{-1})$, and $J = \{J(\omega, \omega_m)\}$ and $K = \{K(\omega, \omega_m)\}$ are $M \times M$ matrices of $\mathcal{J}$ and $\mathcal{K}$ evaluated at the Fourier frequencies, respectively. Throughout this paper, $r$, $i$, and $d$ denote coefficients for real components of $\Theta$, imaginary components of $\Theta$, and the logarithm of the diagonal components of $\Psi^{-1}$, respectively.

As a numerical property of the kernels $\mathcal{J}$ and $\mathcal{K}$ provide a simple form for the roughness of the components such that $P(\Re\{\Theta_{jk}\}) = \theta_{rjk}^2 J_{rjk}$, $P(\Im\{\Theta_{jk}\}) = \theta_{rjk} K_{rjk}$, and $P(\log\Psi_{jj}) = \theta_{djj} J_{djj}$ [17]. Consequently, the prior distributions on the coefficients $z_{rjk}$, $z_{ijk}$, $z_{djj}$ induce a prior on the roughness of the components with $P(\Re\{\Theta_{jk}\}) \sim \lambda_{rjk}^2 \lambda_{rjk}^M$, $P(\Im\{\Theta_{jk}\}) \sim \lambda_{rjk}^2 \lambda_{rjk}^M$, and $P(\log\Psi_{jj}) \sim \lambda_{djj}^2 \lambda_{djj}^M$, where $\lambda_{rjk}^M$ and $\lambda_{djj}^M$ are chi-squared random variables with $M$ degrees of freedom. The smoothing parameters, $\lambda_{rjk}^2$, $\lambda_{ijk}^2$, and $\lambda_{djj}^2$, control the roughness of a Cholesky component such that, as the smoothing parameter is shrunk toward zero, the component tends toward a constant function of frequency with probability 1.

### 2.2.3 Low-rank approximation

The smoothing spline model described in the previous section contains a large number of coefficients, i.e. the number of coefficients is equal to the number of Fourier frequencies. To facilitate the compution and maintain the intuitive interpretation of the prior distribution regularizing roughness as measured through $\mathcal{P}$, we consider the basis formed from the scaled eigenvectors of $\mathcal{J}$ or $\mathcal{K}$ [37, 29, 32]. In particular, the low-rank approximation is formulated equivalently as the Bayesian smoothing spline model at each Fourier frequencies. Denote $\Theta_{jk} = \Theta_{jk}(\omega_1), \ldots, \Theta_{jk}(\omega_M)'$ and $\Psi_{jj}^{-1} = \Psi_{jj}^{-1}(\omega_1), \ldots, \Psi_{jj}^{-1}(\omega_M)'$. Further, let $J = V_J D_J V_J^T$ and $K = V_K D_K V_K^T$ be the spectral decomposition of $\mathcal{J}$ and $\mathcal{K}$, respectively, $Q_J = V_J D^{1/2}_J$, $c_J = D^{1/2}_J J^T z_J$, $Q_K = V_K D^{1/2}_K$, and $c_J = D^{1/2}_K K^T z_J$. The eigenvectors contained in the columns of $V_J$ and $V_K$ are in increasing order of roughness and the eigenvalues contained in the diagonal matrix $D_J$ and $D_K$ decay rapidly [13]. Thus, the Cholesky components can be accurately modeled through the first several columns to provide a low-rank approximation. Let $Q_J$ and $Q_K$ be the $M \times S_J$ and $M \times S_K$ matrices of the first $S_J < M$ and $S_K < M$ eigenvectors, respectively. The selection of $S_J$ and $S_K$ provides a compromise between low-rank computational feasibility and loss of flexibility relative to the full-rank model. We select $S_J = S_K = 10$ in our simulation studies and real data analysis, which accounts for at least 97.975% of the total variance of the full smoothing spline when $T \leq 10^4$ [29]. Consequently, the Cholesky components can be modeled as

$$(4) \quad \Re\{\Theta_{jk}\} = L_J a_{rjk} + Q_J a_{rjk}, \quad j > k = 1, \ldots, p - 1,$$

$$(5) \quad \Im\{\Theta_{jk}\} = Q_K c_{ijk}, \quad j > k = 1, \ldots, p - 1,$$

$$(6) \quad \log\{\Psi_{jj}^{-1}\} = L_J a_{djj} + Q_J c_{djj}, \quad j = 1, \ldots, p,$$

where $L_J$ is a $M \times 2$ matrix with the first column are ones and the second column are $(\omega_1, \ldots, \omega_M)$. We further assume $c_{ijk} \sim N(0, \lambda_{ijk}^2 I_{S_J})$, $a_{djj} \sim N(0, \lambda_{djj}^2 I_{S_J})$, $a_{rjk} \sim N(0, \lambda_{rjk}^2 I_{S_K})$, $a_{rjk} \sim N(0, \sigma_1 I_2)$, and $a_{rjk} \sim N(0, \sigma_2 I_2)$.

### 3. Replicated Time Series

Consider a collection of strictly stationary time series $X_{n,t}$, $t = 1, \ldots, T$, with real-valued static variables, $\nu_n$, observed from $n = 1, \ldots, N$ independent subjects. Without loss of generality, we assume that $\nu_n$ is scaled to take values within $[0, 1]$. The primary goal of this article is to assess the association between the frequency patterns of the time series and the covariates. In this section, we use a tensor-product model for the conditional copula spectral density matrix that extends the Bayesian spline model described in Section 2 for a single time series to account for dependencies on both frequency and outcome.

#### 3.1 The conditional copula spectral density kernel

Recall from (1) that $\gamma_k^l(\tau_1, \tau_2)$ is the lag-$k$ cross-covariance for a single strictly stationary time series. To quantify the association between the copula spectral density of the time series $X_{n,t}$ and the outcome $\nu_n$, we define the conditional copula spectral density kernel, in a similar manner to the conditional power spectrum defined in [29], as

$$f_{\Theta_{\nu_1}, \Theta_{\nu_2}}(\nu, \omega) = \sum_{k \in \mathbb{Z}} \gamma^k_1(\tau_1, \tau_2 | \nu_n = \nu) \exp(-2\pi i\omega k),$$

where $\omega \in \mathbb{R}$, $(\tau_1, \tau_2) \in [0, 1]^2$, $\gamma^k_1(\tau_1, \tau_2 | \nu_n = \nu) = \nu) = \text{Cov}(I\{U_{n,t} \leq \tau_1\}, I\{U_{n,t-k} \leq \tau_2\} | \nu_n = \nu)$, $U_{n,t} := F(X_{n,t})$, and $I\{\cdot\}$ is the indicator function. Similar to the formulation in Section 2, for a set of $\tau = \{\tau_1, \ldots, \tau_p\} \in [0, 1]$, the conditional copula spectral density matrix is defined as

$$f(\nu, \omega) = \left(f_{\Theta_{\nu_1}, \Theta_{\nu_2}}(\nu, \omega)\right)_{j,k=1}^p.$$
As with the copula spectral density matrix of a single time series, \( f(\nu, \omega) \) is a \( p \times p \) positive definite and Hermitian matrix that varies as a function of frequency \( \omega \) and covariate \( \nu \) and characterizes how frequency patterns of the time series vary as a function of the covariate. For fixed \( \nu \), \( f(\nu, \cdot) \) is a periodic and Hermitian function of frequency.

Parallel to the formulation in Section 2.2.2, we consider the modified Cholesky decomposition of the conditional copula spectral density matrix that varies as a function of frequency and characterizes how frequency patterns of the time series vary as a function of the covariate. For fixed \( \nu \), \( f(\nu, \cdot) \) is a periodic and Hermitian function of frequency.

For each of \( p^2 \) Cholesky components, we use the tensor product model of [29] that decomposes the bivariate functions into products of univariate functions of \( \nu \) and \( \omega \) (see Section 3.3).

### 3.2 Bayesian spline for covariates

To aid presentation, we first discuss the Bayesian models for univariate functions of \( \nu \) before formulating the tensor product model of Cholesky components. Since the domain of the scaled outcome values is \([0, 1]\), we consider a restricted spline model where the linear term contains an intercept and slope, and the nonlinear part is defined through the reproducing kernel of the seminorm of \( P \) over restricted smooth functions \( H \). Equivalently, the reproducing kernel Hilbert space can be decomposed as \( 1 \oplus (\nu - 0.5) \oplus H \), where the reproducing kernel for \( H \) is

\[
H(\nu_1, \nu_2) = \kappa_2(\nu_1)\kappa_2(\nu_2) - \kappa_4(\nu_1 - \nu_2), \quad \nu_1, \nu_2 \in [0, 1],
\]

\( \kappa_2(\nu) = |\kappa_2^2(\nu) - 1|/2 \), and \( \kappa_4(\nu) = |\kappa_4^2(\nu) - 2 + 7/240|/24 \) are the scaled Bernoulli polynomials. Again, we use a low-rank approximation of Bayesian smoothing spline model for a function of covariates based on the first \( S_H \) eigenvectors of \( H \),

\[
L_H a + Q_H b.
\]

where \( L_H \) is a \( N \times 2 \) matrix with the first column as ones and the second column as \((\nu_1, \cdots, \nu_N)\)', \( Q_H \) is an \( N \times S_H \) matrix of the first \( S_H \) columns of \( V_H D_H^{1/2} \), \( H = V_H D_H V_H^\prime \) is the spectral decomposition of the \( N \times N \) matrix \( H = \{H(\nu_1, \nu_2)\}, b \sim N(0, \lambda^2 I_{S_H}) \), and \( a \sim N(0, \sigma_a I_2) \).

### 3.3 Bayesian tensor product spline model

Follow [29], we concatenate the Cholesky components across frequency and outcome to define the \( N \times M \)-vectors

\[
\Theta_{jk} = \left\{ (\Theta_{jk}(\nu_1, \omega_1), \cdots, \Theta_{jk}(\nu_1, \omega_M))', \cdots, (\Theta_{jk}(\nu_N, \omega_1), \cdots, \Theta_{jk}(\nu_N, \omega_M))' \right\}.
\]

for \( j > k = 1, \cdots, p - 1 \), and

\[
\Psi_{jj}^{-1} = \left\{ (\Psi_{jj}^{-1}(\nu_1, \omega_1), \cdots, \Psi_{jj}^{-1}(\nu_1, \omega_M))', \cdots, (\Psi_{jj}^{-1}(\nu_N, \omega_1), \cdots, \Psi_{jj}^{-1}(\nu_N, \omega_M))' \right\},
\]

for \( j = 1, \cdots, p \). Combine the Bayesian spline model for functions of frequency defined in (4), (5), and (6) and the model for covariates defined in (8), the tensor product model for the real and imaginary part of \( \theta_{jk} \) and \( \log \psi_{jj}^{-1} \) can be expressed as

\[
\mathbb{R}\{\theta_{jk}\} = \{L_H \otimes L_j \} a_{jk} + \{Q_H \otimes L_j \} b_{jk} + \{L_H \otimes Q_j \} c_{rkj} + \{Q_H \otimes Q_j \} d_{rkj},
\]

\[
\mathbb{I}\{\theta_{jk}\} = \{L_H \otimes Q_K \} c_{rkj} + \{Q_H \otimes Q_K \} d_{rkj},
\]

\[
\log \psi_{jj}^{-1} = \{L_H \otimes L_j \} a_{djj} + \{Q_H \otimes L_j \} b_{djj} + \{L_H \otimes Q_K \} c_{djj} + \{Q_H \otimes Q_K \} d_{djj},
\]

where \( \otimes \) denotes the Kronecker product. This model decomposes the modified Cholesky decomposition defined in (7) as a bivariate function into products of univariate functions of \( \nu \) and \( \omega \). Parameters are coefficients for functions that are products of linear functions of both \( \nu \) and \( \omega \), \( b \) are coefficients for functions that are products of linear functions of \( \omega \) and nonlinear functions of \( \nu \), \( c \) are coefficients for functions that are products of nonlinear functions of \( \omega \) and linear functions of \( \nu \), and \( d \) are coefficients for functions that are products of nonlinear functions of \( \nu \) and \( \omega \). In principle, other Bayesian approaches, such as those in [6, 33], can be adapted to estimate the conditional copula spectral matrix. These adaptive Bayesian methods partition the covariate space into an unknown but finite number of groups in which the number and location of covariate partition points are random and adaptively estimated using reversible-jump Markov chain Monte Carlo (RJMCMC) techniques. However, the RJMCMC could be computationally intensive for a large dataset and may encounter a low acceptance rate problem. Moreover, many replicated time series, such as the stride interval time series, often evolve slowly across covariates, which can be naturally modeled by the proposed smoothing spline ANOVA model that assumes the spectra are smooth in both frequencies and covariates.

Let \( \eta_{rjk} = (\eta_{rjk}^1, \eta_{rjk}^2, \cdots, \eta_{rjk}^{2p}), \eta_{ijk} = (\eta_{ijk}^1, \eta_{ijk}^2, \cdots, \eta_{ijk}^p) \), and \( \eta_{dij} = (\eta_{dij}^1, \eta_{dij}^2, \cdots, \eta_{dij}^p) \), conditional on the smoothing parameters, we assume diagonal Gaussian smoothing priors such that \( \eta_{rjk} \sim N(0, D_{rjk}), \eta_{ijk} \sim N(0, D_{ij}), \) and \( \eta_{dij} \sim N(0, D_{dij}) \), where

\[
D_{rjk} = \text{diag}(\sigma^2_{jk} V_1', \cdots, \sigma^2_{jk} V_p'), \quad D_{ij} = \text{diag}(\sigma^2_{ij} V_1', \cdots, \sigma^2_{ij} V_p'), \quad D_{dij} = \text{diag}(\sigma^2_{dij} V_1', \cdots, \sigma^2_{dij} V_p'),
\]

where \( \lambda_{brjk}, \lambda_{crjk}, \lambda_{brij}, \lambda_{crij}, \lambda_{bdij}, \lambda_{ddij}, j = k = 1, \cdots, p - 1 \), \( \lambda_{bdij}, \lambda_{ddij}, j = 1, \cdots, p \), are the square roots of the smoothing parameters, together, controlling the roughness as a function of frequency and covariate. Prior distributions on the smoothing parameters are assumed to be independent Half-I(\( \gamma, G \)) random variables with pdf \( p(x) \propto [1 + (x/G)^2/\gamma]^{-(\gamma + 1)/2}, x > 0 \), where the hyperparameters \( \gamma \) and \( G \) are assumed known. It has been shown in [15]
that the Half-t prior on the scale parameters is preferred over the uniform and inverse gamma priors. The larger the value of $G$, the less informative the prior, and we set $G$ to a large fixed value. We found inferences are insensitive to the choice of $G$, with $G = 10^6$ giving indistinguishable results in both simulations and the analysis of the gait maturation study. Moreover, inferences are insensitive to the value of $\gamma$, and thus we fix $\gamma = 2$. Computationally, it is convenient to utilize the following scale mixture representation [41]: $(X^2 | g) \sim IG(\gamma/2, \gamma/g)$, $g \sim IG(1/2, 1/G^2)$, where $IG(a, b)$ is the inverse Gamma distribution with pdf $p(x) \propto x^{-(a+1)} \exp(-bx)$, $x > 0$. The hyperparameter $\sigma^2$, which is the variance of the coefficients of the linear terms, is assumed to be a known large value.

3.4 Sampling scheme

As defined in (2), we let the DFT of the $n$th clipped time series at frequency $\omega_m$ be

$$Y_{\tau,n,m} = T^{-1/2} \sum_{t=1}^{T} I(\tilde{F}_p^T(X_{\tau,t}) \leq \tau) \exp(-2\pi i \omega_m t)$$

$$= T^{-1/2} \sum_{t=1}^{T} I(R_{p,t} \leq \tau) \exp(-2\pi i \omega_m t),$$

for $n = 1, \cdots, N$, where $\tilde{F}_p^T(X_{\tau,t}) = T^{-1} \sum_{l=1}^{T} I(X_{\tau, l} \leq x)$ is the empirical marginal distribution function, $R_{p,t}$ is the rank of $X_{\tau,t}$ among $\{X_{\tau,1}, \cdots, X_{\tau,T}\}$. Conditional on $\nu_n$, $Y_{\tau,n,m}$ are approximately independent mean-zero complex Gaussian random variables with large $T$. Let $Y_{nm} = (Y_{\tau,n,m}, \cdots, Y_{r,n,m})^T$; the conditional log likelihood function is

$$L(Y | f) \approx - \sum_{n=1}^{N} \sum_{m=1}^{M} \left\{ \log |f(\nu_n, \omega_m)| + Y_{nm}f^{-1}(\nu_n, \omega_m)Y_{nm} \right\}.$$

Based on the conditional log-likelihood, we develop an efficient MCMC sampling scheme to sample from the joint posterior distribution of the coefficients $\eta$'s and smoothing parameters $\lambda$'s. We denote

$$Q_L = \left( L_H \otimes Q_K : Q_H \otimes Q_K \right),$$

$$Q_r = \left( L_H \otimes L_J : Q_H \otimes Q_J : L_H \otimes Q_J : Q_H \otimes Q_J \right),$$

and $Q_d = Q_r$. Then, (9), (10), and (11) can be expressed respectively

$$\mathbb{R}\{\eta_{jk}\} = Q_r \eta_{jk}, \ \mathbb{R}\{\mu_{jk}\} = Q_L \eta_{jk}, \ \text{and} \ \log \psi_{jk}^{-1} = Q_d \eta_{djk}.$$

Without loss of generality, we describe the sampling scheme for the case $\tau = \{\tau_1, \tau_2, \tau_3\}$ so that $p = 3$. After initializing all the parameters, each iteration of the sampling scheme consists of the following steps:

1. Draw the coefficients $\eta_{rjk}$ for $k < j = 1, 2, 3$ from the conditional posterior distributions

$$\eta_{21} \sim N(\mu_{21}, \Sigma_{21})$$

$$\eta_{31} \sim N(\mu_{31}, \Sigma_{31})$$

$$\eta_{32} \sim N(\mu_{32}, \Sigma_{32}),$$

and the coefficients $\eta_{ij}$ for $k < j = 1, 2, 3$ from the conditional posterior distributions

$$\eta_{i1} \sim N(\mu_{i1}, \Sigma_{i1})$$

$$\eta_{i2} \sim N(\mu_{i2}, \Sigma_{i2}),$$

and then update $\theta_{21} = Q_r \eta_{21} + \sqrt{-1} i Q_l \eta_{21}, \ \theta_{31} = Q_r \eta_{31} + \sqrt{-1} i Q_l \eta_{31}, \ \theta_{32} = Q_r \eta_{32} + \sqrt{-1} i Q_l \eta_{32}$. The exact forms of the conditional means, $\mu_{rjk}$ and $\mu_{ijk}$ and that of conditional covariances, $\Sigma_{rjk}$ and $\Sigma_{ijk}$ for $k < j = 1, \cdots, 3$, are given in the Appendix.

2. Sample $\eta_{dij}$ for $j = 1, 2, 3$, independently through Metropolis–Hastings algorithm from the conditional posterior distribution

$$\log p(\eta_{dij} | Q_d, \rho_{jnm}) \propto \sum_{n=1}^{N} \sum_{m=1}^{M} \left\{ q_{dnm} \eta_{dij} - \exp(q_{dnm} \eta_{dij}) \right\} - \frac{1}{2} \eta_{dij} Q_d^{-1} \eta_{dij},$$

where $q_{dnm}$ is the row of $Q_d$ corresponding to the $n$th subject and $m$th frequency, $\rho_{jnm}$ is a vector with components $\rho_{jnm}$ depending on the DFT and on other parameters held fixed (its exact form is given in the Appendix). More specifically, we draw $\eta_{dij} \sim N(\eta_{dij}, \Sigma_{dij})$, where $\eta_{dij}$ is the maximizer of (12) and $\Sigma_{dij}$ is the inverse of the Hessian matrix at $\eta_{dij}$. The move is accepted with probability $\min(1, \Pi)$, where

$$\Pi = \frac{p(\eta_{dij}^{(c)} | Q_d, \rho_{ji\cdot}, D_{dij}) q(\eta_{dij}^{(c)})}{p(\eta_{dij}^{(p)} | Q_d, \rho_{ji\cdot}, D_{dij}) q(\eta_{dij}^{(p)})},$$

$q$ denotes the normal density, and superscripts $p$ and $c$ denote the proposed and the current values, respectively. Finally, we update $\psi_{jk}^{-1} = \exp(Q_d \eta_{djk})$. In our simulation studies and real data application, the acceptance rate of the proposed method is about 35%–55%. We also don’t encounter any circumstances that no updating occurs.

3. Sample the smoothing parameters $\lambda_{rjk}, \lambda_{rjk}$, and $\lambda_{drjk}$, for $k < j = 1, 2, 3$, from

$$\lambda_{rjk}^2 \sim IG((n_k + \gamma)/2, b'_{jk} b_{rjk}/2 + \gamma/g_{rjk}),$$

$$\lambda_{rjk}^2 \sim IG((n_k + \gamma)/2, c'_{jk} c_{rjk}/2 + \gamma/g_{rjk}),$$

$$\lambda_{drjk}^2 \sim IG((n_d + \gamma)/2, d_{rjk} d_{rjk}/2 + \gamma/g_{drjk}),$$
respectively. Then, update the hyperparameters such that

\[ g_{brjk} \sim IG((\gamma + 1)/2, \gamma/\tau_{brjk}^2 + 1/G^2), \]
\[ g_{crjk} \sim IG((\gamma + 1)/2, \gamma/\tau_{crjk}^2 + 1/G^2), \]
\[ g_{drjk} \sim IG((\gamma + 1)/2, \gamma/\tau_{drjk}^2 + 1/G^2). \]

It should be noted that the smoothing parameters for the imaginary components of \( \Theta \) and the diagonal components can be similarly drawn from inverse gamma distributions.

4. SIMULATIONS

In this section, we evaluate the proposed method through simulated datasets. Three processes, including a conditional AR(2), a conditional QAR(1), and a conditional ARCH(1), are considered. Due to lack of simple closed-form expressions for the true copula spectral densities, we obtain the true copula spectral matrix, \( f(\nu, \omega) \) through numerical simulations [26]. In particular, for each \( n \) (10000) independent replications, of length 1000, of the strictly stationary approximation \( X_{n,t} \) time series were simulated. The actual conditional copula spectra were obtained by computing independent copies of the corresponding CR-periodograms and averaging them across replicates. For each combination of the quantile levels, \( \tau = 0.1, 0.5, 0.9 \), there different numbers of replicated time series \( N = 30, 50, 100 \), and the fixed time series length \( T = 300 \), we compute the root mean squared error (RMSE) of the estimated copula spectral matrix \( f(\nu, \omega) \). The RMSE is obtained by averaging squared errors across the equally-spaced grid of covariate values and frequencies such that

\[
\sqrt{\frac{1}{N(M + 1)} \sum_{n=1}^{N} \sum_{m=0}^{M} \left( \hat{f}_{\nu_1, \nu_2}(\nu_n, \omega_k) - f_{\nu_1, \nu_2}(\nu_n, \omega_k) \right)^2},
\]

where \( M = [(T - 1)/2] \). For each of the 100 repetitions, the proposed method was fit using \( S_I = S_J = SK = 10 \) basis functions, hyperparameters \( G = \sigma^2_\alpha = 10^2 \) and 2000 iterations of the MCMC algorithm with a burn-in of 1000. We also compare the proposed method to two other methods, including the method that smooths the CR-periodograms across frequencies, denoted as SmoothCRP [26]; and the method that further smoothes the estimates of smoothCRP across the replicates by computing the local average of \( B \) neighboring estimates, denoted as RepSmoothCRP-B. In our simulations, we consider \( B = 5, 10 \).

4.1 Conditional AR(2) process

The conditional AR(2) process is simulated as follows:

\[
X_{n,t} = 1.8 \cos\{1.5 - \cos(2\nu_n)\}X_{n,t-1} - 0.81X_{n,t-1} + \epsilon_{n,t},
\]

for \( n = 1, \ldots, N, t = 1, \ldots, T \), where \( \nu_n = n/N \), and \( \epsilon_{n,t} \) is i.i.d. Cauchy noise. Since the Cauchy distribution has an infinite variance, the second-order conditional spectral density does not exist. The conditional copula spectral density, however, provides informative and comprehensive information about the process. For \( N = 100 \), the true conditional copula power spectral matrix and their respective estimates based on the proposed method are presented in Figure 2, in which the lower triangular parts correspond to the real parts of the conditional copula spectral matrix, while the upper triangular parts correspond to the imaginary parts. It should be noted that the significant peak around zero frequency appearing in the diagonals that are associated with extreme quantiles, i.e. \( \tau_1, \tau_2 = 0.1 \) and \( \tau_1, \tau_2 = 0.9 \), indicating persistence in tail events-a phenomenon that cannot be captured traditional analyses. The imaginary parts of the spectra are not close to zero, indicating the non-Gaussian AR(2) process is time-irreversible.

Table 1 displays the mean and standard deviation of RMSEs for estimating components of the conditional copula spectral matrix. The proposed method outperforms SmoothCRP and RepSmoothCRP in all components. The performance of the proposed method and RepSmoothCRP improve as \( N \) increases, while the RMSEs of SmoothCRP stay constant regardless of the number of replicates \( N \). We believe this is because SmoothCRP only smooths across frequencies, and thus does not borrow neighboring information across the replicates to obtain more accurate estimates.

4.2 Conditional QAR(2) and ARCH(1) processes

We consider two additional processes that are temporally uncorrelated. Although with different dynamics, they exhibit the same flat second-order covariance-based spectral density. The copula spectral density kernels, however, provides much richer information about the dynamics of those two processes [26]. First, we generate the conditional QAR(2) process as

\[
X_{n,t} = \{(1.9 U_t - 0.95)\nu_n + (-1.9 U_t + 0.95)(1 - \nu_n)\}X_{n,t-1} + 0.1\Phi^{-1}(U_t),
\]

for \( n = 1, \ldots, N; t = 1, \ldots, T \), where \( \nu_n = n/N \), \( U_t \) is i.i.d. uniform random variable from 0 to 1, and \( \Phi \) denotes the distribution function of the standard normal distribution. Second, we consider the conditional ARCH(1) process as

\[
X_{n,t} = \{1/1.9 + (0.4 + 0.5\nu_n)X_{n,t-1}^2\}^{1/2}\epsilon_{n,t},
\]

for \( n = 1, \ldots, N; t = 1, \ldots, T \), where \( \nu_n = n/N \), and \( \epsilon_{n,t} \) is standard normal white noise.

Tables 2 and 3 present the mean and standard deviation of RMSEs for estimating components of the conditional copula spectral matrices of processes (14) and (15), respectively. Again, RMSEs of the proposed method are significantly smaller than those of the other methods in most.
Figure 2. Rows 1–3: the components of the true conditional copula spectral densities of covariate-dependent AR(2) process (13); Rows 4–6: the corresponding estimated components of the conditional copula spectral densities by the proposed method when $N = 100$. 

Z. Li
Table 1. Estimation accuracy for process (13): reported as mean (standard deviation) of root mean squared error (RMSE) \times 10^2 of components of the conditional copula spectral matrix estimates obtained via the proposed method, SmoothCRP, and RepSmoothCRP. The smallest mean squared error for each component within each setting is highlighted in bold.

<table>
<thead>
<tr>
<th>N</th>
<th>Method</th>
<th>$f_{\tau_{1q},\tau_{1q}}$</th>
<th>$f_{\tau_{2q},\tau_{2q}}$</th>
<th>$f_{\tau_{3q},\tau_{3q}}$</th>
<th>$\Re f_{\tau_{1q},\tau_{1q}}$</th>
<th>$\Re f_{\tau_{2q},\tau_{2q}}$</th>
<th>$\Re f_{\tau_{3q},\tau_{3q}}$</th>
<th>$\Im f_{\tau_{1q},\tau_{1q}}$</th>
<th>$\Im f_{\tau_{2q},\tau_{2q}}$</th>
<th>$\Im f_{\tau_{3q},\tau_{3q}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(0.008)</td>
<td>(0.009)</td>
<td>(0.010)</td>
<td>(0.008)</td>
<td>(0.009)</td>
<td>(0.010)</td>
<td>(0.010)</td>
<td>(0.008)</td>
<td>(0.009)</td>
</tr>
<tr>
<td>30</td>
<td>Proposed</td>
<td>0.865</td>
<td>4.305</td>
<td>1.393</td>
<td>0.815</td>
<td>1.494</td>
<td>0.685</td>
<td>0.047</td>
<td>0.663</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SmoothCRP</td>
<td>1.81</td>
<td>7.70</td>
<td>1.17</td>
<td>0.04</td>
<td>0.005</td>
<td>0.003</td>
<td>0.21</td>
<td>0.681</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RepSmoothCRP-5</td>
<td>1.93</td>
<td>4.33</td>
<td>1.19</td>
<td>0.04</td>
<td>0.005</td>
<td>0.003</td>
<td>0.21</td>
<td>0.681</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RepSmoothCRP-10</td>
<td>1.29</td>
<td>5.50</td>
<td>1.29</td>
<td>0.04</td>
<td>0.005</td>
<td>0.003</td>
<td>0.21</td>
<td>0.681</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>Proposed</td>
<td>0.787</td>
<td>3.998</td>
<td>0.794</td>
<td>0.771</td>
<td>1.439</td>
<td>0.657</td>
<td>0.035</td>
<td>0.658</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SmoothCRP</td>
<td>1.82</td>
<td>7.77</td>
<td>1.17</td>
<td>0.17</td>
<td>0.679</td>
<td>0.04</td>
<td>0.20</td>
<td>0.683</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RepSmoothCRP-5</td>
<td>1.57</td>
<td>4.75</td>
<td>1.15</td>
<td>0.15</td>
<td>0.657</td>
<td>0.04</td>
<td>0.20</td>
<td>0.683</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RepSmoothCRP-10</td>
<td>1.22</td>
<td>5.09</td>
<td>1.22</td>
<td>0.15</td>
<td>0.657</td>
<td>0.04</td>
<td>0.20</td>
<td>0.683</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>Proposed</td>
<td>0.763</td>
<td>3.603</td>
<td>0.679</td>
<td>1.115</td>
<td>0.754</td>
<td>1.229</td>
<td>0.654</td>
<td>0.029</td>
<td>0.650</td>
</tr>
<tr>
<td></td>
<td>SmoothCRP</td>
<td>1.82</td>
<td>7.74</td>
<td>1.17</td>
<td>0.17</td>
<td>0.679</td>
<td>0.04</td>
<td>0.20</td>
<td>0.683</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RepSmoothCRP-5</td>
<td>1.53</td>
<td>4.74</td>
<td>1.15</td>
<td>0.15</td>
<td>0.657</td>
<td>0.04</td>
<td>0.20</td>
<td>0.683</td>
<td></td>
</tr>
<tr>
<td></td>
<td>RepSmoothCRP-10</td>
<td>1.15</td>
<td>4.72</td>
<td>1.15</td>
<td>0.15</td>
<td>0.657</td>
<td>0.04</td>
<td>0.20</td>
<td>0.683</td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Estimation accuracy for process (14): reported as mean (standard deviation) of root mean squared error (RMSE) \times 10^2 of components of the conditional copula spectral matrix estimates obtained via the proposed method, SmoothCRP, and RepSmoothCRP. The smallest mean squared error for each component within each setting is highlighted in bold.

<table>
<thead>
<tr>
<th>N</th>
<th>Method</th>
<th>$f_{\tau_{1q},\tau_{1q}}$</th>
<th>$f_{\tau_{2q},\tau_{2q}}$</th>
<th>$f_{\tau_{3q},\tau_{3q}}$</th>
<th>$\Re f_{\tau_{1q},\tau_{1q}}$</th>
<th>$\Re f_{\tau_{2q},\tau_{2q}}$</th>
<th>$\Re f_{\tau_{3q},\tau_{3q}}$</th>
<th>$\Im f_{\tau_{1q},\tau_{1q}}$</th>
<th>$\Im f_{\tau_{2q},\tau_{2q}}$</th>
<th>$\Im f_{\tau_{3q},\tau_{3q}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(0.008)</td>
<td>(0.009)</td>
<td>(0.010)</td>
<td>(0.008)</td>
<td>(0.009)</td>
<td>(0.010)</td>
<td>(0.010)</td>
<td>(0.008)</td>
<td>(0.009)</td>
</tr>
<tr>
<td>30</td>
<td>Proposed</td>
<td>0.128</td>
<td>0.222</td>
<td>0.116</td>
<td>0.136</td>
<td>0.056</td>
<td>0.134</td>
<td>0.184</td>
<td>0.147</td>
<td>0.186</td>
</tr>
<tr>
<td></td>
<td>SmoothCRP</td>
<td>0.244</td>
<td>0.94</td>
<td>0.24</td>
<td>0.30</td>
<td>1.169</td>
<td>0.306</td>
<td>0.276</td>
<td>0.172</td>
<td>0.273</td>
</tr>
<tr>
<td></td>
<td>RepSmoothCRP-5</td>
<td>0.133</td>
<td>0.31</td>
<td>0.13</td>
<td>0.18</td>
<td>0.074</td>
<td>0.138</td>
<td>0.277</td>
<td>0.174</td>
<td>0.223</td>
</tr>
<tr>
<td></td>
<td>RepSmoothCRP-10</td>
<td>0.126</td>
<td>0.21</td>
<td>0.12</td>
<td>0.14</td>
<td>0.067</td>
<td>0.129</td>
<td>0.205</td>
<td>0.152</td>
<td>0.194</td>
</tr>
<tr>
<td>50</td>
<td>Proposed</td>
<td>0.115</td>
<td>0.171</td>
<td>0.106</td>
<td>0.120</td>
<td>0.046</td>
<td>0.119</td>
<td>0.174</td>
<td>0.140</td>
<td>0.175</td>
</tr>
<tr>
<td></td>
<td>SmoothCRP</td>
<td>0.244</td>
<td>0.69</td>
<td>0.25</td>
<td>0.30</td>
<td>0.171</td>
<td>0.305</td>
<td>0.277</td>
<td>0.173</td>
<td>0.275</td>
</tr>
<tr>
<td></td>
<td>RepSmoothCRP-5</td>
<td>0.128</td>
<td>0.30</td>
<td>0.11</td>
<td>0.13</td>
<td>0.077</td>
<td>0.136</td>
<td>0.276</td>
<td>0.168</td>
<td>0.202</td>
</tr>
<tr>
<td></td>
<td>RepSmoothCRP-10</td>
<td>0.118</td>
<td>0.24</td>
<td>0.11</td>
<td>0.12</td>
<td>0.055</td>
<td>0.120</td>
<td>0.193</td>
<td>0.146</td>
<td>0.193</td>
</tr>
<tr>
<td>100</td>
<td>Proposed</td>
<td>0.104</td>
<td>0.128</td>
<td>0.098</td>
<td>0.109</td>
<td>0.034</td>
<td>0.109</td>
<td>0.163</td>
<td>0.134</td>
<td>0.163</td>
</tr>
<tr>
<td></td>
<td>SmoothCRP</td>
<td>0.245</td>
<td>0.69</td>
<td>0.24</td>
<td>0.30</td>
<td>0.171</td>
<td>0.305</td>
<td>0.274</td>
<td>0.173</td>
<td>0.274</td>
</tr>
<tr>
<td></td>
<td>RepSmoothCRP-5</td>
<td>0.119</td>
<td>0.30</td>
<td>0.10</td>
<td>0.13</td>
<td>0.077</td>
<td>0.136</td>
<td>0.266</td>
<td>0.163</td>
<td>0.193</td>
</tr>
<tr>
<td></td>
<td>RepSmoothCRP-10</td>
<td>0.108</td>
<td>0.23</td>
<td>0.10</td>
<td>0.11</td>
<td>0.054</td>
<td>0.110</td>
<td>0.164</td>
<td>0.135</td>
<td>0.164</td>
</tr>
</tbody>
</table>
Table 3. Estimation accuracy for process (15): reported as mean (standard deviation) of root mean squared error (RMSE) $\times 10^2$ of components of the conditional copula spectral matrix estimates obtained via the proposed method, SmoothCRP, and RepSmoothCRP. The smallest mean squared error for each component within each setting is highlighted in bold.

<table>
<thead>
<tr>
<th>N</th>
<th>Method</th>
<th>$f_{\tau_1,q_1}$</th>
<th>$f_{\tau_2,q_1}$</th>
<th>$f_{\tau_3,q_1}$</th>
<th>$R_{f_{\tau_1,q_1}}$</th>
<th>$R_{f_{\tau_2,q_1}}$</th>
<th>$R_{f_{\tau_3,q_1}}$</th>
<th>$\delta f_{\tau_1,q_1}$</th>
<th>$\delta f_{\tau_2,q_1}$</th>
<th>$\delta f_{\tau_3,q_1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>Proposed</td>
<td>0.206</td>
<td>0.211</td>
<td>0.218</td>
<td>0.104</td>
<td>0.199</td>
<td>0.095</td>
<td>0.031</td>
<td>0.021</td>
<td>0.029</td>
</tr>
<tr>
<td></td>
<td>SmoothCRP</td>
<td>0.294</td>
<td>0.270</td>
<td>0.293</td>
<td>0.312</td>
<td>0.318</td>
<td>0.315</td>
<td>0.263</td>
<td>0.187</td>
<td>0.205</td>
</tr>
<tr>
<td></td>
<td>RepSmoothCRP-5</td>
<td>0.231</td>
<td>0.301</td>
<td>0.234</td>
<td>0.140</td>
<td>0.139</td>
<td>0.077</td>
<td>0.098</td>
<td>0.083</td>
<td>0.119</td>
</tr>
<tr>
<td></td>
<td>RepSmoothCRP-10</td>
<td>0.219</td>
<td>0.215</td>
<td>0.221</td>
<td>0.108</td>
<td>0.123</td>
<td>0.100</td>
<td>0.083</td>
<td>0.058</td>
<td>0.085</td>
</tr>
<tr>
<td>50</td>
<td>Proposed</td>
<td>0.198</td>
<td>0.161</td>
<td>0.212</td>
<td>0.079</td>
<td>0.196</td>
<td>0.073</td>
<td>0.025</td>
<td>0.017</td>
<td>0.022</td>
</tr>
<tr>
<td></td>
<td>SmoothCRP</td>
<td>0.293</td>
<td>0.677</td>
<td>0.296</td>
<td>0.315</td>
<td>0.217</td>
<td>0.315</td>
<td>0.265</td>
<td>0.189</td>
<td>0.264</td>
</tr>
<tr>
<td></td>
<td>RepSmoothCRP-5</td>
<td>0.232</td>
<td>0.303</td>
<td>0.233</td>
<td>0.141</td>
<td>0.139</td>
<td>0.076</td>
<td>0.118</td>
<td>0.084</td>
<td>0.118</td>
</tr>
<tr>
<td></td>
<td>RepSmoothCRP-10</td>
<td>0.218</td>
<td>0.213</td>
<td>0.219</td>
<td>0.109</td>
<td>0.101</td>
<td>0.115</td>
<td>0.088</td>
<td>0.057</td>
<td>0.091</td>
</tr>
<tr>
<td>100</td>
<td>Proposed</td>
<td>0.192</td>
<td>0.115</td>
<td>0.208</td>
<td>0.057</td>
<td>0.191</td>
<td>0.054</td>
<td>0.018</td>
<td>0.011</td>
<td>0.015</td>
</tr>
<tr>
<td></td>
<td>SmoothCRP</td>
<td>0.295</td>
<td>0.676</td>
<td>0.294</td>
<td>0.315</td>
<td>0.217</td>
<td>0.314</td>
<td>0.265</td>
<td>0.189</td>
<td>0.264</td>
</tr>
<tr>
<td></td>
<td>RepSmoothCRP-5</td>
<td>0.229</td>
<td>0.302</td>
<td>0.230</td>
<td>0.132</td>
<td>0.100</td>
<td>0.142</td>
<td>0.119</td>
<td>0.084</td>
<td>0.118</td>
</tr>
<tr>
<td></td>
<td>RepSmoothCRP-10</td>
<td>0.216</td>
<td>0.211</td>
<td>0.218</td>
<td>0.100</td>
<td>0.096</td>
<td>0.101</td>
<td>0.085</td>
<td>0.059</td>
<td>0.083</td>
</tr>
</tbody>
</table>

5. APPLICATION TO THE GAIT MATURATION STUDY

Patterns of gait variability provide insight into how neurological conditions affect the systems that regulate walking [19]. Consequently, stride interval series are often collected to better understand connections between walking patterns and clinical outcomes or demographic information. In this section, the proposed method is applied to study the association between stride interval dynamics and age, which provides a tool for understanding the development of neuromuscular control in children across ages.

We consider a gait maturation study, where the data contains stride interval time series from 50 healthy children (equal numbers of boys and girls) ranging in age from 40 months to 163 months [21]. For each subject, two force-sensitive switches were placed inside the subject’s right shoe: one underneath the heel of the foot and the other underneath the ball of the foot. The output of these foot switches, which provides a measure of the force applied to the floor, was sampled at 300 Hz and stored in a small recorder. These recorded signals were analyzed to determine the initial contact time (heel strike) of each stride throughout the walk, and, hence, the stride time (the time from one heel strike to the next heel strike of the same foot) time series [20]. To eliminate any start-up or ending effects and to allow the subject to become familiar with the walking track, the first 60 seconds and the last 5 seconds of each time series were not included. To make sure that each of the time series has the same length $T$, a total of 350 strides for each subject are considered. Due to possible pauses (stride time $\geq$ 2 second and the 5 seconds before and after any pauses), stride interval time series are subject to outliers (see Figure 1, especially the first two subjects). These outliers present challenges to the spectral analysis as shown in Figure 3, where the sample autocorrelations show that the stride interval time series of the first and second subjects are temporally uncorrelated. Hence, the estimated classical autocovariance-based power spectra [38] cannot detect any significant frequency patterns. In addition, the norm QQ plots presented in Figure 3 indicate that two of these stride time series are heavy-tailed. Although methods can be used to remove outliers, they are subjective and the way to implement them could affect the results significantly. On the other hand, the proposed method, which is based on the copula spectral density kernel, can provide objectively robust spectral analysis of these signals.

The proposed method was fit using $S_H = S_J = S_K = 10$ basis functions, with hyperparameters $G = \sigma_\alpha = 10^5$.  

90 Z. Li
and for 2000 iterations of the MCMC algorithm with a burn-in of 1000. Again, three quantile levels are considered \( \tau = 0.1, 0.5, 0.9 \). In Section 5.1, we present analysis of the estimated conditional copula spectral density as frequency-outcome surfaces. In Section 5.2, we explore the association between the copula spectral density and age through frequency-collapsed band. Convergence assessment of the proposed method is provided in the Appendix.

### 5.1 The conditional copula spectral density

Figure 4 displays components of the estimated conditional copula spectral surfaces. It is clear that the imaginary parts of the spectra are flat across ages, indicating the stride time series are time-reversible. The spectra \( f_{q_{0.5}, q_{0.5}}(\nu, \omega) \), \( \Re f_{q_{0.5}, q_{0.5}}(\nu, \omega) \), and \( \Re f_{q_{0.9}, q_{0.5}}(\nu, \omega) \) show dominant extreme low-frequency components and some fluctuations in high frequency. A closer look reveals that variability in the extremely low frequencies (less than 0.05 Hz) of the real and diagonal parts increases as age increases. This suggests that children with an older age tend to have a locomotor system that possesses memory, such that the change from one stride to the next displays a subtle temporal structure that has been associated with long-range [21]. In addition, the spectra at extreme quantiles, namely, \( f_{q_{0.9}, q_{0.5}}(\nu, \omega) \), \( f_{q_{0.5}, q_{0.1}}(\nu, \omega) \), and \( \Re f_{q_{0.9}, q_{0.1}}(\nu, \omega) \), also have peak or power at low frequencies, but their patterns are relatively constant across ages. This suggests that the stride time series is long-range dependent in general, which is consistent with the findings in the gait variability literature.

#### 5.2 The conditional frequency-collapsed band

Another useful feature of the proposed method in practice is the ability to conduct estimation and inference on any analytic function of the covariate-dependent power spectrum, such as the frequency band summary measure. To better illustrate our findings, we consider two frequency bands: the high-frequency band (HF, 0.25–0.50 Hz) and the low-frequency band (LF, 0.05–0.25 Hz) that are considered in gait variability literature [21]. By computing the ratio of the LF and HF, an index of the frequency walking balance of the spectra is obtained. More specifically, the ratio can be calculated by

\[
\frac{f_{LF/HF}(\nu)}{f_{HF}(\nu)} = \frac{\int_{0.15}^{1.05} f_{q_{1}, q_{2}}(\nu, \omega) d\omega}{\int_{0.50}^{0.25} f_{q_{1}, q_{2}}(\nu, \omega) d\omega},
\]

Figure 5 presents the estimated high-frequency band, low-frequency band, and the LF/HF ratio along with 95%
credible intervals. Our finding suggests that, for all quantiles, the low-frequency bands are relatively constant across age, the high-frequency band increases as age increases, and the ratio of the LF and HF decreases as age increases. This suggests that the gait of the younger children is more nonstationary since a large low-to-high ratio is indicative of walking nonstationarity [21]. Further, it is known that in children of 3 years of age, gait becomes relatively mature; however, there were debates on whether the dynamics of walking change beyond this age. Our results suggest that there are changes in the low-to-high ratio beyond 3 years (36 months), which support the findings that changes in the development of neuromuscular control and locomotor function continue well beyond age 3 [2, 35].

6. DISCUSSION

This article proposes a novel nonparametric method for the simultaneous and automatic analysis of the association between the copula spectral densities of replicated time series and covariates. To the best of our knowledge, the proposed method is the first approach to the analysis of the copula spectral density kernels of replicated time series. The proposed method is robust in the sense that it can handle time series with outliers and heavy-tails. The approach can also provide more important information, such as kurtosis and time-irreversibility that are not shown in the classical spectral analysis. However, the proposed method is not without limitations, which are provided here along with related future extensions to conclude this article. First, the proposed approach considers the impact of a single, fixed covariate on copula spectral densities, which can be extended to consider multiple fixed or time-varying covariates. However, such an extension will lead to a dramatic increase in computational complexity and will likely require some form of simultaneous dimension reduction or efficient approximation of the true nature of the association. Second, the proposed method considers replicated univariate time series. However, in many other applications, such as EEG and fMRI, multivariate or high-dimensional time series are observed. The method of estimating multivariate copula spectral density developed by [1] and the method for high-dimensional time series spectral analysis proposed by [34] provide the possibilities to extend our framework to multivariate or high-dimensional time series. Third, the time series can also vary in time, resulting in nonstationary replicated time series. Future work will
aim to extend the proposed framework to efficiently model the time- and covariate-varying copula spectral density. One possible solution could be through quantile spectral analysis for locally stationary time series [3]. Lastly, other basis functions could also be used. For example, the Demmler-Reinsch bases for periodic even and odd smoothing spline could be used for modeling the Cholesky components of a single time series [37, 43, 32], and cubic penalized splines can be adopted for modeling covariates \( \nu \) [17]. In addition, Bayesian wavelet shrinkage [9] could be used if near line spectra are expected or B-spline bases with knots selected to focus attention within certain frequency bands could be used if it is scientifically known a priori that power is contained within certain frequencies.

**APPENDIX**

Details of the sampling scheme

Assuming \( p = 3 \), we provide more details about the sampling scheme outlined in Section 3.4. Recall that \( q_{nm}, q_{1nm} \) or \( q_{2nm} \) are the rows of \( Q_r, Q_i, \) and \( Q_d \), respectively. The corresponding DFT value is denoted by \( Y_{r,nm}, Y_{i,nm}, \) and \( Y_{d,nm}. \) The \( jkth \) element of \( \Theta(\nu_m, \omega_m) \) defined in Section 3.1 is expressed as

\[
\theta_{jknm} = q_{rnm}\eta_{rjk} + \sqrt{-1}q_{inm}\eta_{ijk}, \quad k > j = 1, \ldots, p - 1,
\]

and \( \theta_{jknm} = q_{rnm}\eta_{rjk} - \sqrt{-1}q_{inm}\eta_{ijk}, \quad k < j = 1, \ldots, p - 1, \)

The diagonal elements of \( \Psi^{-1}(\nu, \omega) \) can be expressed as

\[
\psi^{-1}_{jknm} = \exp(q_{dnm}\eta_{dk}), \quad j = 1, \ldots, p.
\]

To aid presentation, the superscript for iteration number is suppressed and all derived distributions are conditional on the current values of all other parameters.

**Drawing the coefficients of basis functions**

The conditional posterior distributions of \( \eta_{cjk}, c = r, i, j > k = 1, \ldots, p - 1 \), is multivariate normal, \( N(\mu_{cjk}, \Sigma_{cjk}) \).

For \( p = 3 \), the exact forms for \( \mu_{r21} \) and \( \mu_{i21} \) are provided as follows:

\[
\Sigma_{r21}^{-1} = 2 \sum_{n=1}^{N} \sum_{m=1}^{M} \psi_{11nm}^{-1}\left|Y_{r,2,nm}\right|^2 q_{rnm} q_{rnm} + D_{r21}^{-1}
\]

\[
\Sigma_{i21}^{-1} = 2 \sum_{n=1}^{N} \sum_{m=1}^{M} \psi_{11nm}^{-1}\left|Y_{i,2,nm}\right|^2 q_{inm} q_{inm} + D_{i21}^{-1}
\]

\[
\Sigma_{r21}^{-1}\mu_{r21} = 2 \sum_{n=1}^{N} \sum_{m=1}^{M} \psi_{11nm}^{-1}\left|Y_{r,2,nm}\right|^2 q_{rnm} q_{rnm} Y_{r,2,nm} - \theta_{11nm} q_{rnm} Y_{r,2,nm} + \text{Re}\left\{Y_{r,2,nm} Y_{r,2,nm}^* q_{rnm} \right\}
\]

\[
\Sigma_{i21}^{-1}\mu_{i21} = 2 \sum_{n=1}^{N} \sum_{m=1}^{M} \psi_{11nm}^{-1}\left|Y_{i,2,nm}\right|^2 q_{inm} q_{inm} Y_{i,2,nm} + \text{Im}\left\{Y_{i,2,nm} Y_{i,2,nm}^* q_{inm} \right\}
\]

**Figure 5.** For \( \tau = 0.1, 0.5, 0.9, \) the estimated conditional LF (top panel), \( \hat{f}^{LF}_{\eta_r,\eta_r}, \) the estimated conditional HF (middle panel), \( \hat{f}^{HF}_{\eta_r,\eta_r}, \) and the LF/HF (bottom panel), \( \hat{f}^{LF/HF}_{\eta_r,\eta_r}, \) as functions of age with 95% pointwise credible intervals.
Figure 6. Gelman–Rubin Convergence diagnostic plots of the gait maturation study.

\[ \theta_{31nm} Y_{2,nm} Y_{2,nm}^* \] q_{innm},

where \psi_{11nm}^{-1} and \theta_{31nm} are evaluated at their current values.

Mean vectors and covariance matrices for \( \eta_{31} \) and \( \eta_{32} \)

\[
\begin{align*}
\Sigma_{r31}^{-1} &= 2 \sum_{n=1}^{N} \sum_{m=1}^{M} \psi_{11nm}^{-1} |Y_{r1,nm}|^2 q_{rmn} q_{rmn}^* + D_{r31}^{-1} \\
\Sigma_{i31}^{-1} &= 2 \sum_{n=1}^{N} \sum_{m=1}^{M} \psi_{11nm}^{-1} |Y_{i1,nm}|^2 q_{innm} q_{innm}^* + D_{i31}^{-1} \\
\Sigma_{r31}^{-1} \mu_{31} &= 2 \sum_{n=1}^{N} \sum_{m=1}^{M} \psi_{11nm}^{-1} \Re \{ Y_{r1,nm} Y_{r1,nm}^* \} q_{rmn} \\
\Sigma_{i31}^{-1} \mu_{31} &= 2 \sum_{n=1}^{N} \sum_{m=1}^{M} \psi_{11nm}^{-1} \Im \{ Y_{i1,nm} Y_{i1,nm}^* \} q_{innm}.
\end{align*}
\]

Given in Section 3.4, the basis function coefficient vectors \( \eta_{d,j} \), \( j = 1, \ldots, p \) are sampled from \( p(\eta_{d,j} | Q_d, \rho_d, D_{d,j}) \). The entries \( \rho_{jnm} \) of \( \rho_d \), for \( j = 1, 2, 3 \), are as follows.

\[
\rho_{1nm} = |Y_{r1,nm}|^2 + |\theta_{21nm} Y_{r2,nm}|^2 + |\theta_{31nm} Y_{r3,nm}|^2 - 2 \Re \{ \theta_{21nm} Y_{r1,nm} Y_{r2,nm} + \theta_{21nm} Y_{r2,nm} Y_{r3,nm} + \theta_{31nm} Y_{r3,nm} Y_{r3,nm} \}. 
\]
\[ \rho_{2nm} = \frac{|Y_{\tau_2,nm}|^2 + |\theta_{12nm}Y_{\tau_3,nm}|^2 - 2|R\{\theta_{12nm}Y_{\tau_3,nm}\}|}{|Y_{\tau_2,nm}|^2}. \]

\[ \rho_{3nm} = |Y_{\tau_3,nm}|^2. \]

The vectors \( \eta_{djj} \), \( j = 1, \ldots, p \), are generated independently via a Metropolis–Hastings step with a multivariate Normal proposal distribution, \( N(\hat{\eta}_{djj}, \hat{\Sigma}_{djj}) \), where

\[ \hat{\eta}_{djj} = \arg \max_{\eta_{djj}} \log p(\eta_{djj} | Q_d, \rho_j, D_{djj}) \]

and

\[ \hat{\Sigma}_{djj} = \left[ -\frac{\partial^2}{\partial \eta_{djj} \partial \eta_{djj}} \log p(\eta_{djj} | Q_d, \rho_j, D_{djj}) \right]^{-1} \eta_{djj} = \eta_{djj}. \]

The gradient and Hessian of \( \log p(\eta_{djj} | Q_d, \rho_j, D_{djj}) \) are given by

\[
\begin{aligned}
\frac{\sum_{n=1}^{N} \sum_{m=1}^{M} \left[ 1 - \rho_{jnm} \exp(q_{dnn}^n \eta_{djj}) \right] q_{dnn} - D_{djj} \eta_{djj}}{N}
\end{aligned}
\]

and

\[
\begin{aligned}
\frac{- \sum_{n=1}^{N} \sum_{m=1}^{M} \rho_{jnm} \exp(q_{dnn}^n \eta_{djj}) q_{dnn} \eta_{djj} - D_{djj}^{-1}}{N}
\end{aligned}
\]

respectively.

### Convergence assessment of data analysis

In this section, we evaluate the convergence of the proposed method applied to the gait maturation study. We monitor the convergence of the basis function coefficients of the Cholesky components by running \( C > 1 \) simultaneous chains of the sampler with over-dispersed initial values. For example, to track the convergence of \( \psi_{djj} \), we compute a summary measure defined by the largest eigenvalue the matrix \( \eta_{djj} \hat{\eta}_{djj} \) across \( C \) chains and apply the Gelman–Rubin convergence assessment to this derived measure (see details in [16]). With \( C = 5 \) chains, Figure 6 displays the Gelman–Rubin convergence assessments, \( \hat{R} \), for each Cholesky components. It shows that sufficient convergence was achieved since all \( \hat{R} \) converge and stay close to 1.

### ACKNOWLEDGEMENTS

The authors thank the referees, guest editor, and co-editors for providing insightful comments that greatly improved the article. This work was funded in part by PSC-CUNY Research Award 63069-0051 and a Eugene M. Lang Junior Faculty Research Fellowship.

Received 4 January 2021

### REFERENCES


**Robust conditional spectral analysis of replicated time series**


Zeda Li
Paul Chook Department of Information System and Statistics
Baruch College
The City University of New York
One Bernard Baruch Way, New York, New York 10010
The United States of America
E-mail address: zeda.li@baruch.cuny.edu