Statistica Sinica Preprint No: SS-2022-0370	
Title	Test of Partial Separability for Multivariate Functional Data
Manuscript ID	SS-2022-0370
URL	http://www.stat.sinica.edu.tw/statistica/
DOI	10.5705/ss.202022.0370
Complete List of Authors	Fangzhi Luo, Wei Zhang and Decai Liang
Corresponding Authors	Decai Liang
E-mails	liangdecai@nankai.edu.cn
Notice: Accepted version subject to English editing.	

TEST OF PARTIAL SEPARABILITY FOR MULTIVARIATE FUNCTIONAL DATA

Fangzhi Luo, Wei Zhang and Decai Liang

Sun Yat-sen University, Peking University and Nankai University

Abstract: For multivariate functional data, it is quite challenging to model the cross-covariance structure which consists of dual aspects of multivariate and functional features. To simplify the cross-covariance analysis, the assumption of partial separability is widely used to decompose the data into an additive form of multivariate random variables and functional components. In this article, we propose hypothesis testing procedures to examine the validity of partial separability. We study the asymptotic properties of the l^2 and l^∞ norm of the test statistic, resulting in a chi-square type mixture test and a high-dimensional test that are suitable to finite- or high-dimensional multivariate functional data with diverse multivariate dependence. We assess the empirical performance of the proposed tests through two simulation studies for multivariate functional data and graphical functional data, followed by two corresponding real examples: multichannel tonnage data and electroencephalography data.

Key words and phrases: Partial Separability, Multivariate Functional Data, Functional Graphical Model, High-dimensional Test.

1. Introduction

The knowledge of the covariance structure plays an important role in functional data analysis (FDA). To characterize the variation of random curves, nonparametric methods are widely used, often coupled with dimension reduction tools such as functional principal component analysis (FPCA), which provides a flexible approach to depict the temporal dependence. Due to rapid developments in data collection techniques, multivariate functional data (MFD) that comprise simultaneous recordings across multiple processes are becoming increasingly available. Typical examples include daily traffic measurements (Chiou et al., 2014), temperature recordings (Berrendero et al., 2011), multichannel profile data (Paynabar et al., 2016) and neuroimaging data (Happ and Greven, 2018; Qiao et al., 2019). Advanced modeling techniques on the cross-covariance function/kernel/operator, which jointly depict the functional features across time and multivariate dependence across processes, are consequently attracting growing attention.

The analysis of cross-covariance structure has been extensively studied in various fields. For example, the modeling of spatio-temporal covariance plays a prominent part in traditional spatio-temporal statistics, and is highly relevant to a set of spatio-temporal problems such as assessment of stationarity or isotropy, prediction or kriging, resampling methods, and Bayesian inference, see monographs such as (Sherman, 2011; Wikle et al., 2019). To ease the complexity of the covariance structure, a heavily used assumption is separability, which

assumes the full covariance is a product of a purely spatial covariance and a purely temporal covariance. Hypothesis tests of the spatio-temporal separability have been studied in a number of articles (e.g. Fuentes, 2006; Li et al., 2007; Simpson et al., 2014), where restrictions on covariance structure are typically imposed by some parametric forms or stationarity. In these years, some non-parametric tests have been proposed under the functional data setup (Aston et al., 2017; Bagchi et al., 2020; Constantinou et al., 2017), shedding new light on the cross-covariance analysis from the perspective of FDA.

Recently some new notions which relax the separability assumption have been developed, including variants of weak separability for two-way functional data (Lynch and Chen, 2018) and spatio-temporal fields (Liang et al., 2022), and partial separability for multivariate functional data (Zapata et al., 2022). Although introduced for distinct data formats, these definitions are closely interconnected, as discussed in Liang et al. (2022) and Zapata et al. (2022). As the concern of this paper is multivariate functional data, we focus on *partial separability*. We highlight that, as shown in Section 2.1, the partial separability assumption simplifies the full cross-covariance by an eigen-decomposition consisting of a series of functional bases $\{\varphi_l(t) : t \in \mathcal{T}\}_{l=1,2,\ldots}$, i.e. the functional principal components (FPC) which represent the functional variation, and a series of $p \times p$ covariance matrices $\{\Sigma_l\}_{l=1,2,\ldots}$ which reflect the multivariate dependence. Here $\mathcal{T}$ is the functional domain, and p is the multivariate dimension which may be possibly infinite. By a projection onto $\{\varphi_l(t)\}$, the multivariate

functional process $\boldsymbol{X}(t) \in \mathbb{R}^p$ can be further expressed using a sequence of uncorrelated FPC scores $\{\boldsymbol{\theta}_l\}$, each of which is a p-variate random vector and can thus be modeled using techniques in conventional multivariate or high-dimensional statistical analysis.

Due to its flexibility and parsimonious reduction, partial separability has been exploited or implicitly assumed in a large number of literatures. Paynabar et al. (2016) utilized the partial separability representation to simplify the change-point model for multichannel functional profiles. Zapata et al. (2022) proposed the partial separability structure to incorporate the partial correlation for functional graphical models, which overcomes the noninvertibility of the covariance operator for infinite-dimensional functional data. For spatially correlated functional data, partial separability, or the equivalent concept called weak separability in Liang et al. (2022), is widely employed to simplify the modeling of space-time interaction, and is frequently coupled with concepts like stationarity or isotropy (Liu et al., 2017; Zhang and Li, 2021). Note that partial separability is also mixed with the concept of multi-dimensional FPCA (Paynabar et al., 2016), or FPCA in literature on multilevel functional data (Di et al., 2009) and functional point processes (Li and Guan, 2014; Xu et al., 2020), which facilitates substantial model/computational reduction for functional data with complex structures. Despite its versatility to diverse covariance structures, we stress that the partial separability, distinct from univariate or multivariate FPCA based on Karhunen-Loève (KL) or multivariate KL Theorem (e.g. Hsing and

Eubank, 2015), may not hold simultaneously in multivariate and functional aspects, thus its validity should be rigorously examined. Take the example of electroencephalography (EEG) signals in Section 5.2, where one particular interest is to identify the graphical network within the multivariate functional process. However, the separable structure is demonstrably violated in this dataset (e.g. Masak et al., 2023), indicating the need for more flexible dimension reduction approaches like partial separability to simplify the cross-covariance structure effectively. Conducting the proposed statistical tests can offer insights into the appropriateness of adopting partial separability, and the test result can significantly impact subsequent model specifications, including the functional graphical model (Zapata et al., 2022) discussed in Section 2.3.

In this article, we develop statistical tests to formally justify the rationality of assuming partial separability for finite- and high-dimensional multivariate functional data. The test statistics are formulated based on the estimated FPC scores resulting from the marginal covariance kernel. To address the challenge that the test statistic using estimated scores has a different distribution from the counterpart using true scores, we carefully derive the null distribution and demonstrate the theoretical gaps in theorems. In particular, we highlight the importance of constructing different tests for typical multivariate functional data or high-dimensional functional data with diverse dependence structures (e.g. the graphical model). Correspondingly, a χ^2 type mixture test is developed for multivariate functional data with fixed p, in terms of the asymptotic distribution and

properly estimated asymptotic covariance. To adapt to potential high dimensional p and sparse covariance structures, we also propose a high-dimensional test based on the infinite norm of statistics and its Gaussian approximation. Thanks to available replicated observations, the proposed tests do not require other structural assumptions, such as stationarity for spatial functional data (Liang et al., 2022). Also, no Gaussian assumption is required. Our numerical studies suggest that the proposed χ^2 type mixture test or high-dimensional test shows better performance respectively for multivariate functional data or (high-dimensional) graphical functional data.

The remainder of this article is organized as follows. We introduce the notion of partial separability in Section 2 and present its application in several research problems. In Section 3, we propose the testing approaches, involving a χ^2 type mixture test in Section 3.1 and a high-dimensional mean test in Section 3.2. We illustrate the empirical performance of the proposed tests by two simulation studies for multivariate functional data and graphical functional data in Section 4, followed by two application examples on multichannel tonnage data and electroencephalography (EEG) data in Section 5. More technical details and numerical results are provided in the Supplementary Material.

2. Multivariate Functional Data and Partial Separability

Let $L^2(\mathcal{T})$ be the space of square-integrable functions on $\mathcal{T} = [0, 1]$, equipped with inner product $\langle f, g\rangle = \int_{\mathcal{T}} f(t)g(t)dt$ and the corresponding norm $\|\cdot\| =$

$\langle\cdot,\cdot\rangle^{1/2}$. Let $(L^2(\mathcal{T}))^p$ be the space of p-dimensional functions with inner product $\langle \boldsymbol{f}, \boldsymbol{g}\rangle_p = \sum_{j=1}^p \int_{\mathcal{T}} f_j(t)g_j(t)dt$ and norm $\|\cdot\|_p = \langle\cdot,\cdot\rangle_p^{1/2}$. Suppose the multivariate functional data (MFD) are $\boldsymbol{X}(t) = (X_1(t), \ldots, X_p(t))^{\mathrm{T}} \in (L^2(\mathcal{T}))^p$, where each $X_j(t) \in L^2(\mathcal{T})$. We assume that $\boldsymbol{X}(t)$ has a continuous mean function $\boldsymbol{\mu}(t) = (\mu_1(t), \ldots, \mu_p(t))^{\mathrm{T}}$ and a continuous covariance function $\boldsymbol{G}(s,t) = \{G_{jk}(s,t)\} \in \mathbb{R}^{p\times p}$, $1 \le j,k \le p$, following the common approach in FDA (e.g. Hsing and Eubank, 2015; Lynch and Chen, 2018), where $\mu_j(t) = \mathbb{E}\{X_j(t)\}$, $G_{jk}(s,t) = \mathrm{cov}\{X_j(s), X_k(t)\} = \mathbb{E}\{X_j(s) - \mu_j(s)\}\{X_k(t) - \mu_k(t)\}$. Note that $\boldsymbol{G}(s,t)$ and $G_{jk}(s,t)$ can also be regarded as the kernel of the covariance operator $\boldsymbol{\mathcal{G}}$ of $\boldsymbol{X}$ and the cross-covariance operator $\mathcal{G}_{jk}$ of $\{X_j, X_k\}$ (Hsing and Eubank, 2015); see more details in the Supplementary Material.

2.1 The concept of partial separability

According to the above definition of (cross) covariance function, $\boldsymbol{G}(s,t)$ consists of a collection of $p\times p$ matrices across $s,t \in \mathcal{T}$, say $\{G_{jk}(s,t)\}$, which is generally complicated due to the dual aspects of multivariate dependence and infinite-dimensional nature of functional data. Following Zapata et al. (2022), we state the definition of partial separability as follows in terms of the cross-covariance kernel.

Definition 1. *A cross-covariance $\boldsymbol{G}(s,t) = \{G_{jk}(s,t)\}_{j,k=1,\ldots p}$ is partially separable if there exist some orthonormal basis $\{\varphi_l(\cdot)\}_{l=1}^{\infty}$ of $L^2(\mathcal{T})$ and a sequence of $p \times p$ positive-definite matrices $\{\boldsymbol{\Sigma}_l = (\sigma_l(j,k))_{j,k=1,\ldots p}\}$ such that for any*

$j, k = 1, \ldots, p$,

$$G_{jk}(s,t) = \sum_{l=1}^{\infty} \sigma_l(j,k)\, \varphi_l(s)\varphi_l(t). \tag{2.1}$$

Based on Definition 1 and similarly to the Karhunen-Loéve (KL) Theorem, we have the following equivalent definition of partial separability for the random process $\boldsymbol{X}(t)$.

Definition 1†. *A multivariate functional process $\boldsymbol{X}(t)$ is partially separable if there exist an orthonormal basis $\{\varphi_l(\cdot)\}_{l=1}^{\infty}$ of $L^2(\mathcal{T})$, such that*

$$\boldsymbol{X}(t) = \boldsymbol{\mu}(t) + \sum_{l=1}^{\infty} \boldsymbol{\theta}_l\, \varphi_l(t) \tag{2.2}$$

holds almost surely in $(L^2(\mathcal{T}))^p$, where $\boldsymbol{\theta}_l = (\langle X_1 - \mu_1, \varphi_l \rangle, \ldots, \langle X_p - \mu_p, \varphi_l \rangle)^{\mathrm{T}}$ is a p-variate random vector, and $cov(\boldsymbol{\theta}_l, \boldsymbol{\theta}_{l'}) = 0$ for $l \neq l'$, i.e., the scores $\{\boldsymbol{\theta}_l\}$ are mutually uncorrelated.

The equivalence of Definition 1 and 1† is apparent by noticing that the expansion (2.1) has no cross-terms across different l due to the uncorrelatedness of FPC scores $\boldsymbol{\theta}_l$. And it can be easily observed that the matrix $\boldsymbol{\Sigma}_l$ in Definition 1 corresponds to the covariance matrix of $\boldsymbol{\theta}_l$. See Theorem 1 of Zapata et al. (2022) for more equivalent definitions. In what follows we mix up the partial separability of $\boldsymbol{X}$ or $\boldsymbol{G}$ without confusion, and the expansions in (2.1) and (2.2) are ordered in terms of the decreasing value of $\mathrm{tr}(\boldsymbol{\Sigma}_l)$ for clarity. We also assume that $\mathbb{E}\boldsymbol{X}(t) = 0$ without loss of generality.

According to equation (2.2), the orthonormal basis $\{\varphi_l(t)\}$ in (2.1) consists of the eigenfunction of $G_{jk}(s,t)$ for any j, k. Besides, as stated by Lynch and Chen

(2018); Zapata et al. (2022), $\{\varphi_l(t)\}$ is unique up to a sign, and $\varphi_l(t) \equiv \phi_l(t)$, where $\{\phi_l(t)\}$ are the eigenfunctions of the marginal kernel

$$H(s,t) = \sum_{j=1}^{p} G_{jj}(s,t). \tag{2.3}$$

Zapata et al. (2022) show that the eigenbasis of $H(s,t)$ is optimal in the sense that it retains the largest amount of total variability among vectors of the form $(\langle X_1, \tilde{\varphi}_l\rangle, \ldots, \langle X_p, \tilde{\varphi}_l\rangle)^{\mathrm{T}}$ for any orthonormal $\{\tilde{\varphi}_l\}$. If $\boldsymbol{X}$ is partially separable, the unique basis $\{\varphi_l\}$ corresponds to this optimal basis, and

$$H(s,t) = \sum_{l=1}^{\infty} \lambda_l \, \varphi_l(s)\varphi_l(t), \tag{2.4}$$

where $\lambda_l = \mathrm{tr}\,(\boldsymbol{\Sigma}_l)$. Equation (2.4) is useful to construct the estimator of φ_l in Section 3.

2.2 Properties of partial separability

For a more comprehensive understanding, we elucidate the relationship between partial separability and some other concepts, including separability, univariate and multivariate FPCA.

Partial separability and separability. According to Zapata et al. (2022), a multivariate functional process $\boldsymbol{X}(t)$ with the cross-covariance $\boldsymbol{G}(s,t) = \{G_{jk}(s,t)\}_{j,k=1,\ldots p}$ is *separable* if there exist a $p \times p$ covariance matrice $\boldsymbol{\Sigma} = (\sigma(j,k))_{j,k=1,\ldots,p}$ and a covariance function $C(s,t)$ with $s,t \in \mathcal{T}$ such that for any $j,k = 1,\ldots,p$,

$$G_{jk}(s,t) = \sigma(j,k)\, C(s,t). \tag{2.5}$$

As shown by (2.5), the separability assumption allows the full cross-covariance function $\boldsymbol{G}$ to be factorized as a product of a (multivariate) covariance $\boldsymbol{\Sigma}$ and a (functional) covariance $C(s,t)$. By contrast, a partial separable cross-covariance (2.1) is characterized by a sequence of covariance matrices $\boldsymbol{\Sigma}_l$, yielding a more flexible yet parsimonious representation. A more comprehensive result for the relationship between separability and partial separability is given in the following Proposition 1, which is a similar result to Proposition 1 and 2 in Liang et al. (2022)

Proposition 1. *(a) A separable process $\boldsymbol{X}(t)$ is partially separable.*
(b) A partially separable $\boldsymbol{X}(t)$ satisfying (2.1) is separable if and only if (i.f.f) there exist a nonnegative decreasing series of ω_l satisfying $\sum_{l=1}^{\infty}\omega_l < \infty$ and a function $\tilde{\sigma}(j,k)$ such that $\sigma_l(j,k) = \omega_l\,\tilde{\sigma}(j,k)$ for any $j,k = 1,\ldots,p$.

Partial Separability and FPCA. We clarify the difference between partial separability and univariate or multivariate FPCA. The univariate FPCA represents each component of $\boldsymbol{X}(t)$, say $X_j(t)$, in terms of a KL expansion, by

$$X_j(t) = \sum_{l=1}^{\infty} \langle X_j, \phi_{lj}\rangle\, \phi_{lj}(t), \tag{2.6}$$

where the FPC scores $\xi_{jl} = \langle X_j, \phi_{lj}\rangle$ are uncorrelated across different l, $\{\phi_{lj}\}$ are the eigenfunctions of covariance kernel $G_{jj}(s,t)$ and vary across different j. In contrast, partial separability (2.2) assumes a common eigenbasis $\{\varphi_l\}$ for different elements of $\boldsymbol{X}(t)$, which is valid when multiple curves exhibit similar patterns and is useful to integrate the information across multiple components.

On the other hand, Chiou et al. (2014) proposed the following multivariate FPCA representation

$$\boldsymbol{X}(t) = \sum_{l=1}^{\infty} \langle \boldsymbol{X}, \boldsymbol{\psi}_l \rangle_p \, \boldsymbol{\psi}_l(t) \tag{2.7}$$

based on a multivariate KL expansion. Here $\boldsymbol{\psi}_l(t) = (\psi_{l1}(t), \ldots, \psi_{lp}(t))^{\mathrm{T}} \in (L^2(\mathcal{T}))^p$ is the eigenfunction of covariance $\boldsymbol{G}(s,t)$, and $\langle \boldsymbol{\psi}_l(t), \boldsymbol{\psi}_{l'}(t) \rangle_p = \delta_{ll'}$, where $\delta_{ll'} = 1$ if $l = l'$ and 0 otherwise. Despite the parsimony of decomposition (2.7), the FPC scores $\langle \boldsymbol{X}, \boldsymbol{\psi}_l \rangle_p$ are scalar and cannot be directly modeled using approaches in multivariate statistical analysis. By contrast, the partial separability model combines the strengths of (2.6) and (2.7) through a sequence of FPCs that depict the common functional variation among all components, and a series of scores that account for the diversity of multivariate dependence.

For a more thorough overview, we list some other concepts of FPCA or separability in existing literatures. Jiang and Wang (2010) developed a covariate-adjusted FPCA by adjusting the mean and/or covariance function through some covariates that are not available for the considered multivariate functional data. The common functional principal component model in Benko et al. (2009) aimed at identifying common FPCs from two populations, whereas the proposed partial separability accounts for the cross-covariance among p different elements/populations where p may tend to infinity. Recently Masak et al. (2023) defined a novel "R-separable" structure based on the partial inner product that generalizes the notion of separability. Unlike partial separability that yields a straightforward representation (2.2) for the multivariate functional process, "R-

separability" only focuses on simplifying the covariance kernel and is therefore less efficient in producing powerful modeling schemes, such as the functional graphical model discussed in Section 2.3. There is also a notable connection between partial separability and the weak separability for two-way functional data proposed by Lynch and Chen (2018). Recall the expansion (2.2) and suppose that the multivariate functional process $\boldsymbol{X}(t)$ be regarded as a two-way functional process $X \in L^2(S \times \mathcal{T})$, where $S = 1, \ldots, p$. For each vector $\boldsymbol{\theta}_l$, one may further consider a projection $\boldsymbol{\theta}_l = \sum_{k=1}^{p} r_{lk}\boldsymbol{\psi}_{lk}$ onto the p-variate basis $\{\boldsymbol{\psi}_{lk}\}_{k=1,\ldots,p}$ based on the conventional multivariate PCA. If $\{\psi_{lk}\}$ is independent of l, i.e. each of $\{\boldsymbol{\theta}_l\}$ has the same eigen basis $\{\psi_k\}$, then the partial separablity model (2.2) would become $\boldsymbol{X}(t) = \boldsymbol{\mu}(t) + \sum_{l=1}^{\infty}\sum_{k=1}^{p} r_{lk}\boldsymbol{\psi}_k\varphi_l(t)$, which corresponds to the weak separability in Lynch and Chen (2018) that is a special case of our definition.

2.3 Application of partial separability

To illustrate the advantage of partial separability, we give a brief review of its application in two typical research topics.

Change-point detection for multichannel profiles. In statistical process control, change-point detection remains a challenging problem for multi-stream profile monitoring. Paynabar et al. (2016) represents these multichannel profiles through a set of vector-valued functional data $\{\boldsymbol{X}_i(t), i = 1, \ldots, n\}$, where $\boldsymbol{X}_i(t) = (X_{i1}(t), \ldots, X_{ip}(t))^{\mathrm{T}}$ is the vector of observed profiles measured

at $t \in \mathcal{T}$. As the multiple profile curves demonstrate similar patterns, the partial separability model (2.2) is assumed, yielding common FPCs for different profiles. Following a multidimensional FPCA procedure, a sequence of projected vectors $\hat{\boldsymbol{\eta}}_{\tau,l} \in \mathbb{R}^p$ are estimated based on the estimated FPC scores $\hat{\boldsymbol{\theta}}_l \in \mathbb{R}^p$ as well as the classical binary segmentation approach, where l is the component index and τ is the potential change-point. Utilizing the estimator $\widehat{\boldsymbol{\Sigma}}_l$ for $\boldsymbol{\Sigma}_l \in \mathbb{R}^{p\times p}$, that is the asymptotic covariance of $\hat{\boldsymbol{\eta}}_{\tau,l}$, and the mutual uncorrelatedness among $\hat{\boldsymbol{\eta}}_{\tau,l}$ due to partial separability, an additive form of Wald statistic is proposed by

$$Q_\tau = \sum_l \hat{\boldsymbol{\eta}}_{\tau,l}^{\mathrm{T}} \, \widehat{\boldsymbol{\Sigma}}_l^{-1} \, \hat{\boldsymbol{\eta}}_{\tau,l}, \tag{2.8}$$

of which the maximum value across $1 \le \tau \le n$ can be used to detect the change point.

Functional graphical models. The traditional Gaussian graphical models have been extended to multivariate functional objects in some recent research (e.g. Qiao et al., 2019; Zapata et al., 2022). Consider a multivariate process $\boldsymbol{X}(t) \in (L^2(\mathcal{T}))^p$ with an undirected graph (V, E), where $V = \{1, \ldots, p\}$ is the node set and $E \subset V^2$ is the edge set. Let $\boldsymbol{X}_{V\setminus\{j,k\}}$ denote the sub-vector of $\boldsymbol{X}(\cdot)$ obtained by removing $X_j(\cdot)$ and $X_k(\cdot)$. A functional Gaussian graphical model represents the conditional dependency of $\boldsymbol{X}(t)$ by the conditional covariance kernels

$$C_{jk}(s,t) = \mathrm{cov}\left\{X_j(s), X_k(t) | \boldsymbol{X}_{V\setminus\{j,k\}}\right\}, \quad j,k \in V, j \neq k,$$

which is related to the graph of V by that $(j,k) \in E$ i.f.f. $C_{jk}(s,t) = 0$ for all

$s, t \in \mathcal{T}$. Zapata et al. (2022) state that, under a partially separable structure (2.1) or (2.2) with the covariance $\mathbf{\Sigma}_l$, the conditional covariance $C_{jk}(s,t)$ yields the following decomposition

$$C_{jk}(s,t) = \sum_{l=1}^{\infty} c_l(j,k)\varphi_l(s)\varphi_l(t) \tag{2.9}$$

where $c_l(j,k) = \sigma_l(j,k) - [\mathbf{\Sigma}_l]_{j,V\setminus\{j,k\}}\{[\mathbf{\Sigma}_l]_{V\setminus\{j,k\},V\setminus\{j,k\}}\}^{-1}[\mathbf{\Sigma}_l]_{V\setminus\{j,k\},k}$. Here $[\cdot]_{V_1,V_2}$ denotes the operation to extract subsets V_1 and V_2 of rows and columns of a matrix. Let $\mathbf{\Omega}_l = \mathbf{\Sigma}_l^{-1}$ be the precision matrix of $\boldsymbol{\theta}_l$, with each element being $\omega_l(j,k)$, $j,k = 1,\ldots,p$. It can be shown that $c_l(j,k) = 0$ i.f.f. $\omega_l(j,k) = 0$, thus $(j,k) \notin E$ i.f.f. $c_l(j,k) = \omega_l(j,k) = 0$ for all l. We highlight that noticing the similarity between (2.1) and (2.9), the conditional dependency can also be characterized by a sequence of precision matrices $\{\mathbf{\Omega}_l\}$ owing to the partial separability, which is useful to overcome the problem of noninvertibility of the covariance operator when $\boldsymbol{X}$ is infinite-dimensional. It should also be noted that this functional graphical model is applicable when p is infinite, and the estimation procedure can be embedded with some high-dimensional statistical tools such as the graphical lasso (Danaher et al., 2014).

It is common knowledge that the graph structure in functional graphical models may be sparse, and it is of great interest when the dimension p is large. This motivates us to develop a substantially different test procedure (a high-dimensional test in Section 3.2) from the traditional multivariate analysis (a χ^2 type mixture test in Section 3.1). One may also consider other types of multivariate functional data with specific covariance structures, e.g. the spatial

correlation that is usually smaller for stations at a closer distance. Correspondingly, the partially separable covariance structure, also called weak separability (Liang et al., 2022) or coregionalization model (Zhang and Li, 2021), is widely employed in semiparametric modeling of spatial functional data. We also apply the proposed tests for spatially correlated functional data, and collect relevant results in Supplementary Material for space economy.

3. Tests of Partial Separability

According to Section 2.1, a p-variate process $\boldsymbol{X}(t)$ can be projected onto a unique orthonormal basis $\varphi_l(\cdot)$ consisting of the eigen-system of the marginal covariance $H(s,t)$, which results in a series of vectors $\boldsymbol{\theta}_l \in \mathbb{R}^p$. Let $\boldsymbol{\theta}_l = (\theta_{l1}, \ldots, \theta_{lp})^{\mathrm{T}}$ and $\boldsymbol{\Sigma}_{ll'} = \mathrm{cov}(\boldsymbol{\theta}_l, \boldsymbol{\theta}_{l'})$. By Definition 1†, testing the partial separability of $\boldsymbol{X}(t)$ is the same as testing the covariance $\boldsymbol{\Sigma}_{ll'} = 0$ for $l \neq l'$, or equivalently,

$$\mathbb{H}_0 : \mathrm{cov}(\theta_{lj}, \theta_{l'k}) = 0 \text{ for all } l \neq l' \text{ and } j, k = 1, \ldots, p. \tag{3.10}$$

If one truncates $\boldsymbol{X}(t)$ in (2.2) with L components, the partial separability test is also to examine if the covariance matrix of $(\boldsymbol{\theta}_1^{\mathrm{T}}, \ldots, \boldsymbol{\theta}_L^{\mathrm{T}})^{\mathrm{T}}$ is block diagonal; see Figure 1 of Zapata et al. (2022) for a graphical illustration.

Suppose we observe $\boldsymbol{X}_i(t) = (X_{i1}(t), \ldots, X_{ip}(t))^{\mathrm{T}}, i = 1, \ldots, n$, which are i.i.d. copies of $\boldsymbol{X}(t)$. Since the sample values of $\{\boldsymbol{\theta}_l\}$ cannot be directly observed, we first estimate the scores by

$$\hat{\boldsymbol{\theta}}_{i,l} = (\hat{\theta}_{i,l1}, \ldots, \hat{\theta}_{i,lp})^{\mathrm{T}} = \left(\langle X_{i1} - \hat{\mu}_1, \hat{\varphi}_l \rangle, \ldots, \langle X_{ip} - \hat{\mu}_p, \hat{\varphi}_l \rangle\right)^{\mathrm{T}}, \tag{3.11}$$

based on the conventional projection approach. Here $\hat{\mu}_j(t) = n^{-1}\sum_{i=1}^n X_{ij}(t)$ is the sample mean function, $\{\hat{\varphi}_l(t)\}$ are eigenfunctions of the estimated marginal covariance

$$\widehat{H}(s,t) = \sum_{j=1}^{p} \widehat{G}_{jj}(s,t), \tag{3.12}$$

where the cross-covariance $G_{jk}(s,t)$ is obtained by

$$\widehat{G}_{jk}(s,t) = \frac{1}{n}\sum_{i=1}^{n} \{X_{ij}(s) - \hat{\mu}_j(s)\}\{X_{ik}(t) - \hat{\mu}_k(t)\} \tag{3.13}$$

for each $j,k = 1,\ldots,p$. It can be easily observed that $\widehat{G}_{jk}(s,t)$, $\widehat{H}(s,t)$, $\hat{\varphi}_l(t)$ and $\hat{\boldsymbol{\theta}}_{i,l}$ are consistent to their population counterparts. In practice, if the functional data are sparsely observed with measurement errors, we can also estimate G_{jk} through the smoothing procedure on pooled data (e.g. Yao et al., 2005; Chiou et al., 2014).

As a common practice in FDA, we consider a truncated process $\boldsymbol{X}(t) = \boldsymbol{\mu}(t) + \sum_{l=1}^{L_n} \boldsymbol{\theta}_l\, \varphi_l(t)$ with a suitable chosen L_n, of which the practical selection is usually tricky and depends on the subsequent analysis. Provided with the eigenvalues $\{\hat{\lambda}_l\}$ of the marginal covariance $\widehat{H}$, one may determine L_n according to the fraction of variance explained (FVE), define as $\text{FVE}(L_n) = (\sum_{l=1}^{L_n}\hat{\lambda}_l)/(\sum_{l=1}^{\infty}\hat{\lambda}_l)$, which can be practically chosen as 80%, 90% or 95% etc. Then, to test the covariance structure of $\boldsymbol{\theta}_l$ and $\boldsymbol{\theta}_{l'}$ for each $1 \le l < l' \le L_n$, a reasonable statistic is

$$T_n(l,j,l',k) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} \hat{\theta}_{i,lj}\hat{\theta}_{i,l'k}, \tag{3.14}$$

across all $j,k = 1,\ldots,p$. Let $T_n^{(l,l')} = (T_n(l,j,l',k) : 1 \le j \le k \le q)^{\mathrm{T}}$ be a

$p(p+1)/2$-vector by omitting the duplicate terms for $j > k$, and $\boldsymbol{T}_n$ be a long vector by stacking $\{T_n^{(l,l')} : 1 \le l < l' \le L_n\}$ with length $q = p(p+1)L_n(L_n-1)/4$. Based on the l^2- and l^∞- norm of $\boldsymbol{T}_n$, we then develop a χ^2 type mixture test in Section 3.1, as well as a high-dimensional test in Section 3.2 to cope with the situation when p or L_n is large.

3.1 Test based on χ^2 type mixture

As stated by Lynch and Chen (2018); Liang et al. (2022), although the estimate scores $\{\hat{\boldsymbol{\theta}}_{i,l}\}$ can be shown to be $\sqrt{n}$-consistent, the test statistic (3.14) has a different distribution from the counterpart using true scores. To provide the asymptotic result of $\boldsymbol{T}_n$, we first state two conditions which are necessary for the main theorem.

Condition 1. $\mathbb{E}\|\boldsymbol{X}\|_p^4 < \infty$.

Condition 2. *The eigenvalues of $H(s,t)$, say $\{\lambda_l\}$, have multiplicity one.*

Condition 1 is a standard moment assumption for the central limit theorem of (cross) covariance operators in Hilbert-Schmidt spaces (Theorem 8.1.2 in Hsing and Eubank, 2015). Following Theorem 9.1.3 in Hsing and Eubank (2015), Conditions 1 and 2 ensure the weak convergence of the estimated eigenfunctions.

Theorem 1. *Assume Conditions 1 and 2 hold and $\boldsymbol{X}$ is partially separable.*

(a) For each $1 \le l < l' \le L_n$ and $1 \le j \le k \le p$, we have

$$T_n(l,j,l',k) = Z_n(l,j,l',k) + o_p(1),$$

where $Z_n(l, j, l', k)$ converges weakly to a mean-zero Gaussian random variable.

(b) $\boldsymbol{T}_n$ is asymptotically jointly Gaussian with mean zero and covariance structure $\boldsymbol{\Theta}$.

The explicit expressions of $Z_n(l, j, l', k)$ and $\boldsymbol{\Theta}$ are given in the Supplementary Material.

To perform the test in terms of the asymptotic joint normality in Theorem 1, a χ^2 test statistic $\boldsymbol{T}_n^T \boldsymbol{\Theta}^{-1} \boldsymbol{T}_n$ may be considered. However, one could easily observe a linear relationship among the diagonal elements in $\boldsymbol{T}_n$, i.e., $\{T_n(l, j, l', j) : j = 1, \ldots, p\}$, by noting

$$\begin{aligned}
\sqrt{n} \sum_{j=1}^{p} T_n(l, j, l', j) &= \sum_{j=1}^{p} \sum_{i=1}^{n} \hat{\theta}_{i,lj} \hat{\theta}_{i,l'j} \\
&= \sum_{j=1}^{p} \sum_{i=1}^{n} \int \{X_{ij}(t) - \hat{\mu}_j(t)\} \hat{\varphi}_l(t) dt \int \{X_{ij}(s) - \hat{\mu}_j(s)\} \hat{\varphi}_{l'}(s) ds \\
&= \sum_{j=1}^{p} \int \int \hat{G}_{jj}(s, t) \hat{\varphi}_l(s) \hat{\varphi}_{l'}(t) ds dt \\
&= \int \int \hat{H}(s, t) \hat{\varphi}_l(s) \hat{\varphi}_{l'}(t) ds dt = 0, \qquad (3.15)
\end{aligned}$$

which implies that $\boldsymbol{\Theta}$ is singular. Instead, we consider the non-normalized statistic constructed by summation of the square of $T_n(l, j, l', k)$, say

$$S_n = \sum_{1 \le l < l' \le L_n} \sum_{1 \le j \le k \le p} T_n(l, j, l', k)^2, \qquad (3.16)$$

following a similar procedure in Lynch and Chen (2018). In light of Theorem 1, one could easily obtain the distribution of S_n under null/alternative hypothesis, as presented in the following corollary.

Corollary 1. *Assume the conditions in Theorem 1 hold.*

(a) If $\boldsymbol{X}$ *is partially separable, then* S_n *follows a* χ^2 *type mixture distribution; more explicitly,* $S_n = \sum_{i=1}^{q} \gamma_i Z_i^2$*, where* $\{\gamma_i\}$ *are the eigenvalues of the* $q \times q$ *matrix* $\boldsymbol{\Theta}$*, and* $\{Z_i\}$ *are i.i.d standard normal random variables.*

(b) If $\boldsymbol{X}$ *is non-partially separable, then* $S_n \to \infty$ *in probability.*

Based on Corollary 1 we could then perform a χ^2 type mixture test. As suggested by Zhang (2013), the exact distribution of a χ^2 type mixture is generally not tractable, especially when q is large. Consequently, the Welch-Satterthwaite approximation is commonly-used and approximates $S_n \sim \beta\chi_d^2$, where β and d are determined by matching the first two moments of S_n. This results in $\beta = tr(\boldsymbol{\Theta}^2)/tr(\boldsymbol{\Theta})$ and $d = \{tr(\boldsymbol{\Theta})\}^2/tr(\boldsymbol{\Theta}^2)$. Similar to Lynch and Chen (2018); Liang et al. (2022), the explicit form of $\boldsymbol{\Theta}$ depends on the cross fourth-order moments of the FPC scores, say $\mathbb{E}(\theta_{lj_1}\theta_{l'k_1}\theta_{lj_2}\theta_{l'k_2})$, which in practice can be estimated via moment estimation based on the estimated scores $\hat{\boldsymbol{\theta}}_{i,l}$. Thus a plug-in estimator of $\boldsymbol{\Theta}$ can be used to approximate the p-values for the proposed χ^2 type mixture test as an upper tail probability of $\beta\chi_d^2$.

3.2 High-dimensional Test

For many problems in multivariate functional data analysis, it would be more interesting to consider partial separability when the dimension p is high, such as the functional graphical model in Qiao et al. (2019); Zapata et al. (2022). In these cases, executing the χ^2 type mixture test in Section 3.1 on $\boldsymbol{T}_n$ may lead

to bad performance due to the curse of dimensionality. For this consideration, we propose a high-dimensional test for partial separability motivated by the distribution and correlation-free test of high-dimensional means proposed in Xue and Yao (2020).

Recall the definition of $\boldsymbol{T}_n$ and consider the counterpart of $T_n^{(l,l')}$ based on the true scores, say

$$Y^{(l,l')} = (\theta_{l1}\theta_{l'1}, \ldots, \theta_{l1}\theta_{l'p}, \ldots, \theta_{lp}\theta_{l'p})^{\mathrm{T}},$$

similarly let $\boldsymbol{Y}$ be a long vector by stacking $\{Y^{(l,l')} : l < l'\}$. Denote $\mu_{\boldsymbol{Y}} = \mathbb{E}\boldsymbol{Y}$, we can reformulate the test (3.10) as

$$H_0 : \mu_{\boldsymbol{Y}} = 0 \quad \text{v.s.} \quad H_1 : \mu_{\boldsymbol{Y}} \neq 0.$$

For a large p, this is actually a high-dimensional one-sample mean test with the dimension of $\boldsymbol{Y}$ being $q = p(p+1)L_n(L_n-1)/4$, identical to that of $\boldsymbol{T}_n$. For an infinite-dimensional functional process $\boldsymbol{X}(t)$, the truncation L_n may also be allowed to diverge, which possibly makes the dimension q even larger. Let $S_n^{\boldsymbol{Y}} = n^{-1/2}\sum_{i=1}^n \boldsymbol{Y}_i$, and denote the q-dimensional vector $S_n^{\boldsymbol{Y}}$ as $\{S_{n,d}^{\boldsymbol{Y}} : d = 1, \ldots, q\}$. As the usual approach in high-dimensional statistics, we consider the l^∞-norm statistic

$$\left\|S_n^{\boldsymbol{Y}}\right\|_\infty = \max_{1\le d\le q} |S_{n,d}^{\boldsymbol{Y}}|. \tag{3.17}$$

Then the test rejects the null hypothesis at a certain significant level α if $\left\|S_n^{\boldsymbol{Y}}\right\|_\infty > c(\alpha)$ for some critical value $c(\alpha)$.

To obtain a proper critical value, Xue and Yao (2020) applied a multiplier bootstrap method for a high-dimensional two-sample test problem, based on the high-dimensional central limit theorem (Chernozhukov et al., 2017). Following their procedure, we define $S_n^{e\boldsymbol{Y}} = n^{-1/2}\sum_{i=1}^n e_i\left(\boldsymbol{Y}_i - \hat{\mu}_{\boldsymbol{Y}}\right)$, where $\{e_i\}_{i=1}^n$ is a set of i.i.d. standard normal random variables independent of the data and $\hat{\mu}_{\boldsymbol{Y}} = n^{-1}\sum_{i=1}^n \boldsymbol{Y}_i$ denotes the sample mean. This multiplier bootstrap procedure is useful to generate resamples of Gaussian approximations to $S_n^{\boldsymbol{Y}}$. To quantify the error of approximation between $\left\|S_n^{\boldsymbol{Y}}\right\|_\infty$ and $\left\|S_n^{e\boldsymbol{Y}}\right\|_\infty$, we require the following condition.

Condition 3. *Denote* $\mathbf{Y}_i = (Y_{i,1},\ldots,Y_{i,q})^{\mathrm{T}}$, $S_n^{\boldsymbol{Y}} = (S_{n,1}^{\boldsymbol{Y}},\ldots,S_{n,q}^{\boldsymbol{Y}})^{\mathrm{T}}$.

(a) There exists a universal constant $b_1 > 0$ *s.t.* $\min_{1\le d\le q}\mathbb{E}\left\{\left(S_{n,d}^{\boldsymbol{Y}}\right)^2\right\} \ge b_1$.

(b) There exists a sequence of constants $B_n \ge 1$ *such that*

$$\max_{1\le d\le q}\sum_{i=1}^n \mathbb{E}\left(\left|Y_{i,d} - \mathbb{E}Y_{i,d}\right|^3\right)/n \le B_n, \quad \max_{1\le d\le q}\sum_{i=1}^n \mathbb{E}\left(\left|Y_{i,d} - \mathbb{E}Y_{i,d}\right|^4\right)/n \le B_n^2.$$

(c) The sequence of constants B_n *in (b) also satisfies*

$$\max_{1\le i\le n}\max_{1\le d\le q}\mathbb{E}\left\{\exp\left(\left|Y_{i,d}\right|/B_n\right)\right\} \le 2.$$

(d) $B_n^2\left\{\log(qn)\right\}^7/n \to 0$ *as* $n\to\infty$.

Condition 3(a) and (b) correspond to the moment properties of the coordinates, while Condition 3(c) is associated with the tail properties. No other restriction on the distribution or correlation of those random vectors is required. More explanations about these conditions can be seen in Xue and Yao (2020). In

particular, Condition 3(d) indicates that the dimension q can grow exponentially in n, provided that B_n is of some appropriate order.

Theorem 2. *Assume Condition 3 holds and $\boldsymbol{X}(t)$ is partially separable. Then with probability one, the Kolmogorov distance between the distributions of $\left\|S_n^{\boldsymbol{Y}}\right\|_\infty$ and $\left\|S_n^{e\boldsymbol{Y}}\right\|_\infty$ satisfies*

$$\sup_{t\geq 0}\left|\,\mathbb{P}\left(\left\|S_n^{\boldsymbol{Y}}\right\|_\infty \leq t\right)-\mathbb{P}_e\left(\left\|S_n^{e\boldsymbol{Y}}\right\|_\infty \leq t\right)\,\right| \lesssim \left\{B_n^2\log^7(qn)/n\right\}^{1/6},$$

where $a_n \lesssim b_n$ means $a_n \leq Cb_n$ up to a universal constant $C>0$ for two sequence of constants a_n and b_n, $\mathbb{P}_e(\cdot)$ denotes the probability with respect to $\{e_i\}_{i=1}^n$ only. Consequently,

$$\sup_{\alpha\in(0,1)}\left|\mathbb{P}\left\{\left\|S_n^{\boldsymbol{Y}}\right\|_\infty \leq c_B(\alpha)\right\}-(1-\alpha)\right| \lesssim \left\{B_n^2\log^7(qn)/n\right\}^{1/6},$$

where $c_B(\alpha)=\inf\left\{t\in\mathbb{R}:\mathbb{P}_e\left(\left\|S_n^{e\boldsymbol{Y}}\right\|_\infty \leq t\right)\geq 1-\alpha\right\}$.

This theorem provides a simpler result on the one-sample mean test than the two-sample test result in Theorem 3 of Xue and Yao (2020), and could also be straightforwardly implied by the one-sample CLT in Chernozhukov et al. (2017), so we omit its proof. To implement the test procedure based on Theorem 2, we first obtain the estimate $\widehat{\boldsymbol{Y}}_i$ and $S_n^{\widehat{\boldsymbol{Y}}}$ through the consistently estimated scores $\hat{\boldsymbol{\theta}}_{i,l}$ defined in (3.11), and then perform the multiplier bootstrap based on the estimate of $S_n^{e\widehat{\boldsymbol{Y}}}$. The detailed algorithm is displayed in Algorithm 1.

Algorithm 1 Algorithm for High-dimensional Test

1: **For** $i = 1 \ldots, n$ **do**

2: Compute $\widehat{\boldsymbol{Y}}_i$ by stacking $\{\widehat{\boldsymbol{Y}}_i^{(l,l')} : 1 \le l < l' \le L_n\}$, where $\widehat{\boldsymbol{Y}}_i^{(l,l')} = (\hat{\theta}_{i,l1}\hat{\theta}_{i,l'1}, \ldots, \hat{\theta}_{i,lp}\hat{\theta}_{i,l'p})^{\mathrm{T}}$, i.e, $\widehat{\boldsymbol{Y}}_i = \{(\widehat{\boldsymbol{Y}}_i^{(1,2)})^{\mathrm{T}}, \ldots, (\widehat{\boldsymbol{Y}}_i^{(1,L_n)})^{\mathrm{T}}, \ldots, (\widehat{\boldsymbol{Y}}_i^{(L_n-1,L_n)})^{\mathrm{T}}\}^{\mathrm{T}}$.

3: Compute $S_n^{\widehat{\boldsymbol{Y}}} = n^{-1/2}\sum_{i=1}^n \widehat{\boldsymbol{Y}}_i$. Note from (3.14) that $S_n^{\widehat{\boldsymbol{Y}}} = \boldsymbol{T}_n$.

4: **For** $b = 1, \ldots, B$ **do**

5: Generate n samples $e_1, \ldots, e_n$ from standard normal distribution.

6: Compute statistics $\mathcal{S}_n^e = \left\| S_n^{e\widehat{\boldsymbol{Y}}} \right\|_\infty$ while keeping $\widehat{\boldsymbol{Y}} = \{\widehat{\boldsymbol{Y}}_1, \ldots, \widehat{\boldsymbol{Y}}_n\}$ fixed, where $S_n^{e\widehat{\boldsymbol{Y}}} = n^{-1/2}\sum_{i=1}^n e_i\left(\widehat{\boldsymbol{Y}}_i - \hat{\mu}_{\widehat{\boldsymbol{Y}}}\right)$ and $\hat{\mu}_{\widehat{\boldsymbol{Y}}}$ is the sample mean of $\widehat{\boldsymbol{Y}}$.

7: Obtain the approximated critical value $\hat{c}_B(\alpha)$ by the $100(1-\alpha)$-th quantile of $\left\{\mathcal{S}_n^{e1}, \ldots, \mathcal{S}_n^{eB}\right\}$ with $\alpha \in (0,1)$.

8: Reject $\mathbb{H}_0$ if $\left\| S_n^{\widehat{\boldsymbol{Y}}} \right\|_\infty > \hat{c}_B(\alpha)$.

4. Simulation Studies

We conduct numerical experiments to investigate the test performance for typical multivariate functional data and graphical functional data with potential high dimension. For both simulation studies, we set the mean function as 0 and the time domain as $[0, 1]$ with 30 equally spaced grids. The fraction of variance explained (FVE) is set to be 80% or 90%. The empirical size or power is evaluated via 1000 or 200 runs respectively, at the 0.05 significant value.

4.1 Multivariate Functional Data

Following the expansion (2.2), for $i = 1, \ldots, n$ we generate a mean-zero random vector $\boldsymbol{\theta}_i = (\boldsymbol{\theta}_{i,1}^{\mathrm{T}}, \ldots, \boldsymbol{\theta}_{i,L}^{\mathrm{T}})^{\mathrm{T}} \in \mathbb{R}^{pL}$ with $\boldsymbol{\theta}_{i,l} = (\theta_{i,l1}, \ldots, \theta_{i,lp})^{\mathrm{T}} \in \mathbb{R}^{p}$, and then the multivariate functional data $\boldsymbol{X}_i(t)$ by

$$\boldsymbol{X}_i(t) = \sum_{l=1}^{L} \boldsymbol{\theta}_{i,l}\, \varphi_l(t), \tag{4.18}$$

where $\{\varphi_l(t)\}$ are the first four non-constant Fourier basis functions. We let $p = 4$, $L = 4$ and $n = 200$. The covariance matrix $\boldsymbol{\Sigma} \in \mathbb{R}^{pL \times pL}$ of $\boldsymbol{\theta}_i$ is determined according to the following cases:

(i) Partially separable case: $\boldsymbol{\Sigma}_{ps} = \mathrm{diag}(\boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_L)$ with $\boldsymbol{\Sigma}_l = (\sigma_l(j,k)) \in \mathbb{R}^{p \times p}$ and $\sigma_l(j,k) = \mathrm{cov}\{\theta_{i,lj}, \theta_{i,lk}\} = l^{-2}\, r^{|j-k|}$ (Paynabar et al., 2016).

(ii) Non-partially separable case: $\boldsymbol{\Sigma}_{nps}$ with the diagonal block being $\boldsymbol{\Sigma}_{l,l} = \boldsymbol{\Sigma}_l$ in (i) and the non-zero off-diagonal blocks $\boldsymbol{\Sigma}_{1,2} = \boldsymbol{\Sigma}_{2,1} = \rho(\boldsymbol{\Sigma}_1^* + \boldsymbol{\Sigma}_2^*)$ where $\boldsymbol{\Sigma}_l^* = \boldsymbol{\Sigma}_l - \mathrm{diag}(\boldsymbol{\Sigma}_l)$ and $\rho > 0$, which means that the covariance of $\boldsymbol{\theta}_{i,1}$ and $\boldsymbol{\theta}_{i,2}$ is non-zero.

One can see that the constructed $\boldsymbol{\Sigma}_{nps}$ is positive-definite and would be equal to $\boldsymbol{\Sigma}_{ps}$ when $\rho = 0$, yielding the null hypothesis. In addition, the parameter r quantifies the correlation between multivariate objects. The departure from partial separability would be stronger, i.e. the non-zero elements in $\boldsymbol{\Sigma}_{1,2}$ would be larger as ρ or r increases. Provided with the covariance $\boldsymbol{\Sigma}$, the vector $\boldsymbol{\theta}_i$ is simulated from $N(0, \boldsymbol{\Sigma})$ or the multivariate t distribution. For the latter case,

Table 1: Empirical rejection rates(%) of χ^2 type mixture test (χ^2) and high-dimensional test (HD) for typical multivariate functional data.

		$\rho = 0$ ($\mathbb{H}_0$)		ρ=0.3		ρ=0.5		ρ=0.7	
r	FVE	χ^2	HD	χ^2	HD	χ^2	HD	χ^2	HD
					Normal				
0.2	80%	4.5	3.8	15.0	10.5	51.0	38.5	85.5	78.0
	90%	5.2	3.8	14.5	10.5	48.5	37.5	82.5	78.0
0.3	80%	4.4	4.1	42.5	26.0	89.5	83.5	100	100
	90%	4.9	4.1	37.5	26.0	90.0	83.5	100	100
0.4	80%	4.8	3.7	70.0	52.5	100	99.0	100	100
	90%	4.6	3.7	67.5	52.5	100	99.0	100	100
					Multivariate t				
0.2	80%	4.5	2.4	10.5	7.0	38.5	22.5	72.0	54.5
	90%	4.3	2.4	12.5	6.5	36.0	22.0	72.0	53.5
0.3	80%	4.6	2.5	33.0	20.5	79.0	64.5	98.5	92.5
	90%	4.3	2.4	30.0	19.5	76.0	64.0	99.0	92.0
0.4	80%	4.6	2.5	56.5	35.0	98.0	90.5	100	100
	90%	4.0	2.2	53.5	34.0	98.0	90.0	100	100

we first generate a vector $\boldsymbol{x}$ from $N(0, \boldsymbol{\Sigma})$ and a χ^2 random variable u with a degree of freedom v that is independent of $\boldsymbol{x}$. Following Lynch and Chen (2018), we use $\boldsymbol{x}/\{u/(v-2)\}^{1/2}$ as our multivariate t vector such that its covariance is $\boldsymbol{\Sigma}$, and take $v = 6$ in our simulations. Besides the partial separability test,

it is noteworthy that the covariance in case (i) is also separable according to Proposition 1. Therefore, it would be interesting to conduct the separability test on the simulated data; the corresponding results are included in Section S4 of the Supplementary Material.

Table 1 displays the empirical rejection rate results with varying ρ and r under the scenario of normal or multivariate t distribution. We can see that the χ^2 type mixture test is able to control type I error well under different cases, while the high-dimensional test tends to be undersized especially for the multivariate t case. Besides, the empirical power rises rapidly as ρ grows from 0.3 to 0.7 or r from 0.2 to 0.4, and the χ^2 test performs uniformly more powerful than the high-dimensional test, indicating the superiority of χ^2 type mixture test for typical multivariate functional data. Table 1 also shows that the tests achieve stable Type I errors under different FVE, and perform slightly more powerful when FVE= 80%. Note that the selected L_n has an average of 2 or 3 respectively for FVE= 80% or 90% in 200 trials. As the non-separable components occur in the off-diagonal block $\mathbf{\Sigma}_{1,2}$ for $\mathbf{\Sigma}_{nps}$, the power is reasonably better when $L_n = 2$ is mostly chosen, i.e., FVE= 80%. Our additional result indicates that the proposed FVE criterion is adaptive to different decay rates and effectively determines the value of L_n; see Section S3.2 of Supplementary Material.

4.2 Graphical Functional Data

We start with the multivariate functional model (4.18) with $L = 11$ and then add the graph structure on $\boldsymbol{\theta}_{i,l}$ by the following procedure. Given the node set $V = \{1, \ldots, p\}$ and a sparsity parameter π, we first randomly generate an edge between any two nodes with probability $2\pi/(p-1)$ so that there are πp edges on average. Based on such a graph $G = (V, E)$, we then generate a sequence of edge sets $\{E_l\}_{l=1,\ldots,L}$ so that $E = \cup_{l=1}^{L} E_l$, and a sequence of precision matrices $\{\Omega_l\}_{l=1,\ldots,L}$ for each E_l; for brevity we omit the details of the generating algorithm which can be found in Zapata et al. (2022). Then we generate a mean-zero pL-variate Gaussian vector $\boldsymbol{\theta}_i$ with the covariance matrix $\boldsymbol{\Sigma}$ as follows:

(i†) $\boldsymbol{\Sigma}_{ps} = \mathrm{diag}(\boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_L)$ with $\boldsymbol{\Sigma}_l = 3l^{-2}\Omega_l^{-1} \in \mathbb{R}^{p\times p}$.

(ii†) $\boldsymbol{\Sigma}_{nps}$ with the diagonal block $\boldsymbol{\Sigma}_l$ in (i†) and the off-diagonal blocks $\boldsymbol{\Sigma}_{1,3} = \boldsymbol{\Sigma}_{3,1} = \rho(\boldsymbol{\Sigma}_1^* + \boldsymbol{\Sigma}_3^*)$, where $\boldsymbol{\Sigma}_l^* = \boldsymbol{\Sigma}_l - \mathrm{diag}(\boldsymbol{\Sigma}_l)$ and $\rho = 0.3$.

To study the test power under a different alternative case, in scenario (ii†) we impose the signal at the 1st and 3rd components, which generalizes the non-partially separable case (ii) in Section 4.1. Besides the correlation parameter ρ, the violation of partial separability is also influenced by the sparsity level π, that is, the number of nonzero elements in Ω_l and $\boldsymbol{\Omega}_{nps} := \boldsymbol{\Sigma}_{nps}^{-1}$ would grow when π becomes larger.

As suggested by Table 2, both tests control the type I error well, whereas the high-dimensional test performs uniformly more powerful than the χ^2 type

Table 2: Empirical rejection rates(%) of χ^2 type mixture test (χ^2) and high-dimensional test (HD) for graphical functional data.

		π=0.05		π=0.1		π=0.2		π=0.3	
p	FVE	χ^2	HD	χ^2	HD	χ^2	HD	χ^2	HD
Size ($\rho = 0$)									
25	80%	5.0	4.1	5.7	4.4	4.8	4.6	4.6	4.4
	90%	5.5	4.1	5.6	4.4	5.0	4.6	4.8	4.4
50	80%	5.9	5.2	4.4	4.3	4.8	4.1	3.8	5.3
	90%	6.1	5.2	4.9	4.3	5.3	4.1	4.2	5.3
100	80%	4.1	3.9	4.4	4.0	4.5	3.9	4.5	4.2
	90%	4.4	3.9	4.5	4.0	5.4	3.9	4.6	4.2
Power ($\rho = 0.3$)									
25	80%	30.0	58.0	30.0	55.5	18.5	25.0	12.0	16.5
	90%	45.5	97.0	52.0	98.0	68.5	98.5	87.0	100
50	80%	16.5	83.0	24.0	71.5	30.0	66.5	42.0	63.0
	90%	18.5	95.0	32.0	98.5	44.5	99.5	65.0	99.5
100	80%	12.5	92.0	21.0	94.5	32.5	94.5	43.5	89.5
	90%	14.0	95.0	22.5	98.5	34.5	99.5	43.5	100

mixture test, especially when π is small (e.g. $\pi = 0.05$) which corresponds to a rather sparse graph. Moreover, as p increases the high-dimensional test has a stable power, by contrast with χ^2 type mixture which becomes less powerful for larger p (especially for FVE=90%). These results demonstrate the superiority of

high-dimensional test for graphical functional data. In addition, we can observe an obvious improvement of power from FVE= 80% to 90%. Since the non-separable components in (ii†) occur in $\mathbf{\Sigma}_{1,3}$, and the selected L_n is 2.6 or 4 on average respectively for FVE= 80% or 90%. As expected, the tests behave more powerful when $L_n \geq 3$ is mostly chosen, i.e., FVE= 90%. This also indicates that the claim of partial separability should be a comprehensive conclusion across different FVE or L_n. In addition, we investigate a more complex simulation setting that mimics the graph structure in the real application (see Section S3.1 of Supplementary Material), and the obtained results are consistent with the above findings.

5. Real Data Examples

5.1 Multichannel Tonnage Profiles

The multichannel tonnage profile data were previously studied for Phase-I monitoring change detection (Lei et al., 2010; Paynabar et al., 2016). This dataset contains a collection of four-channel tonnage profile curves, each of which is recorded by a strain gauge sensor mounted on one column of the forging machine. The data sample used in the analysis includes 151 in-control(IC) profiles under the normal condition, followed by one of five out-of-control (OC) groups with 69 profiles, resulting in $n = 220$ subjects. Following Paynabar et al. (2016), we consider the data of Fault 4 group, which contains the most similar profile curves to the normal condition and thus is difficult to separate. To reduce the

measurement noises, we use B-spline functions with 30 bases to smooth each curve, reducing each profile function to 200 equally spaced time points.

To account for the multivariate dependence within multiple profile curves, Paynabar et al. (2016) propose a change point model which embeds multidimensional FPCA into the binary segmentation procedure. As pointed out in Section 2.3, partial separability is assumed to borrow information from all channels, and plays a central role in the construction of change detection statistics. It is thus necessary to examine the validity of the assumed partial separability structure. Since the multivariate data contains $p = 4$ channel profiles and exhibit typical multivariate functional data patterns, we perform the test based on χ^2 type mixture as suggested in Section 4.1. As the mean function of IC and OC groups are different, the IC and OC profile data are firstly centralized separately by subtracting the corresponding sample mean.

Table 3 displays the test results with different truncation levels and different subgroups of channels. As shown by the first row, the hypothesis of partial separability is rejected for the whole dataset with all 4 channels across all levels of truncation. This suggests that it may not be appropriate to assume a partial separability structure directly on the 4-variate profile curves. To give a more comprehensive analysis, we also conduct the χ^2 type mixture test on the subgroups of channels $(2, 3, 4)$ and $(3, 4)$. It can be seen that for channels $(2, 3, 4)$, the p-values are less than 0.001 when $L_n > 2$, i.e. the FVE is larger than 60%; while the null hypothesis is not rejected for channels $(3, 4)$ across different trun-

Table 3: The p-values of partial separability tests based on χ^2 type mixture with different truncation levels for multichannel tonnage data.

Channels	L_n	2	5	8	11	14
(1,2,3,4)	FVE	52.1%	76.2%	83.9%	89.0%	93.1%
	p-value	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
(2,3,4)	FVE	58.2%	76.5%	85.1%	90.2%	93.7%
	p-value	0.380	< 0.001	< 0.001	< 0.001	< 0.001
(3,4)	FVE	57.0%	76.8%	85.8%	91.0%	94.2%
	p-value	0.800	0.378	0.366	0.356	0.351

cation levels, which supports the validity of partial separability assumption on this subset of 2-variate functional data.

To give some illustration of the partial separability structure, we compare the eigenfunctions resulting from the marginal covariance $\widehat{H}(s,t) = \sum_{j=1}^{p} \widehat{C}_{jj}(s,t)$ defined by (3.12) with the univariate FPCs of each channel, as shown in Figure 1. The univariate FPCs are matched in order by minimizing the integrated square errors with the eigenfunctions of $\widehat{H}$. It can be observed that, despite an overall common pattern for the first eigenfunctions, the disparity becomes evident from the second column of eigenfunctions for channels (1,2,3,4). Specifically, the second eigenfunction of channel 1 (black solid line) has an obvious drift compared to other channels, and the difference becomes more significant for the third eigenfunction. This illustrates the violation of partial separabil-

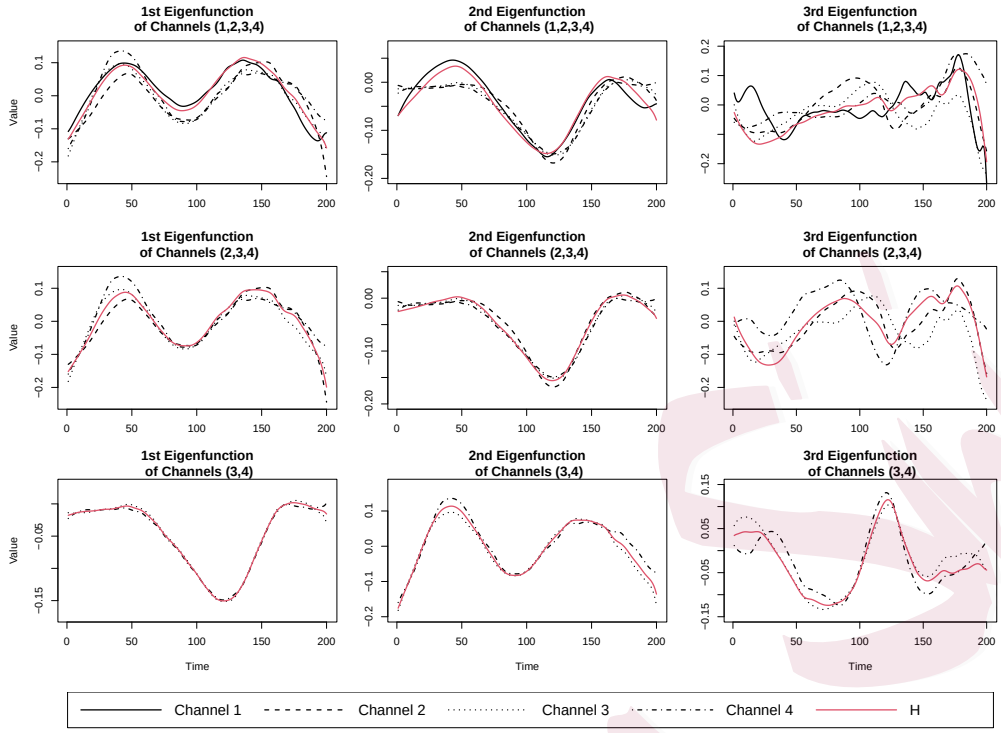

Figure 1: Estimated eigenfunctions of the marginal covariance $\widehat{H}$ (red) and univariate FPCs of each channel for tonnage data.

ity which assumes all profile channels have common eigenfunctions. As for the middle row of channels (2,3,4), the inconsistent variation emerges at the third eigenfunctions, which supports our findings of the non-partial separability when $L_n > 2$. By contrast for channels (3,4), the first three univariate FPCs and eigenfunctions of $\widehat{H}$ all exhibit a similar pattern, which provides evidence for the partially separable assumption on this subgroup.

5.2 EEG Dataset

To investigate the performance of our method for graphical functional data, we apply the proposed partial separability test to the EEG dataset from an alcoholism study (Zhang et al., 1995). This study consists of $n = 122$ subjects with 77 in the alcoholic group and 45 in the control group. For each subject, voltage values are measured from $p = 64$ electrodes, which are recorded at $T = 256$ time points for one second. Hu and Yao (2022) propose a sparse high-dimensional FPCA approach to extract the dynamic features of EEG signals. Qiao et al. (2019) apply a functional graphical model to identify the connectivity between different brain regions of interest. The functional graphical model is further developed by Zapata et al. (2022) under the partial separability assumption, as discussed in Section 2.3, which theoretically overcomes the noninvertibility of infinite-dimensional covariance operator and is demonstrated to perform better for functional data with higher variance explained.

We conduct the same data preprocessing as Qiao et al. (2019) using an α band filtering and consider the average signal of all trials under the single stimulus condition. Then the eigendecomposition is performed on the marginal covariance $\widehat{H}$, which shows that the FVEs for $L_n = 3, 6, 9$ are respectively $58.9\%, 81.5\%, 93.7\%$. As suggested by the simulation results in Section 4.2, we apply the high-dimensional test for the partial separability under these L_n's, resulting in the corresponding p-values as displayed in the first row (PS-T) of Table 4. These results indicate that assuming partial separability with an ade-

Table 4: The p-values of the partial separability test (PS-T) and the separability test (Sep-T) of Aston et al. (2017) with different L_n for EEG data.

L_n	3	6	9
PS-T	0.135	0.121	0.110
Sep-T	< 0.001	< 0.001	< 0.001

quate truncation level could be a justifiable approach for a subsequent functional graphical model in Zapata et al. (2022). We also notice that the p-value for the χ^2 type mixture test is less than 0.05 when L_n is larger than 6, which may be due to a relatively small sample size ($n = 122$), compared with the length of $\boldsymbol{T}_n$ being $q = 31200$ when $L_n = 6$.

One intriguing finding is that the separability assumption is demonstrably inappropriate by the separability test proposed by Aston et al. (2017). The corresponding p-values are shown in the second row of Table 4, while additional details are provided in Section S4.2 of Supplementary Material. This result provides additional evidence that partial separability is more flexible than separability in accommodating the cross-covariance structure of the multivariate functional data.

6. Discussion

The proposed tests of partial separability can significantly influence the subsequent analysis of specific problems. Our further studies on the multichannel

tonnage data in Section 5.1 reveal that misuse of the partial separability assumption can potentially lead to incorrect change-point detection; additional results are included in Supplementary Material. When dealing with sparsely observed functional data, our test may be extended by utilizing smoothed covariance estimators based on pooled data (Yao et al., 2005). However, the distribution of the test statistic may be affected by inflated estimation errors, which require further investigation. Another theoretical challenge arises when the truncation level is allowed to diverge to infinity, which reflects the nonparametric nature of functional data. This would lead to more complicated convergence results for the FPC according to the perturbation theory (e.g. Hsing and Eubank, 2015), necessitating more involved technical developments for the corresponding test.

Supplementary Materials

The Supplementary Material contains the proofs of Theorem 1 and Corollary 1, additional numerical results for spatially correlated functional data, high dimensional functional data, decay rates of eigenvalues and the relationship between partial separability and separability, as well as subsequent analysis for partial separability and separability tests on the specific applications.

Acknowledgements

Wei Zhang is the co-first author. This research is supported by the National Natural Science Foundation of China Grants No. 12101332, 12231017 and 12292984,

the Key Laboratory of Pure Mathematics and Combinatorics (LPMC) and the Key Laboratory for Medical Data Analysis and Statistical Research of Tianjin (KLMDASR).

References

Aston, J. A., D. Pigoli, S. Tavakoli, et al. (2017). Tests for separability in nonparametric covariance operators of random surfaces. The Annals of Statistics 45(4), 1431–1461.

Bagchi, P., H. Dette, et al. (2020). A test for separability in covariance operators of random surfaces. Annals of Statistics 48(4), 2303–2322.

Benko, M., W. Härdle, and A. Kneip (2009). Common functional principal components. The Annals of Statistics, 1–34.

Berrendero, J. R., A. Justel, and M. Svarc (2011). Principal components for multivariate functional data. Computational Statistics & Data Analysis 55(9), 2619–2634.

Chernozhukov, V., D. Chetverikov, and K. Kato (2017). Central limit theorems and bootstrap in high dimensions. The Annals of Probability 45(4), 2309–2352.

Chiou, J.-M., Y.-T. Chen, and Y.-F. Yang (2014). Multivariate functional principal component analysis: A normalization approach. Statistica Sinica, 1571–1596.

Constantinou, P., P. Kokoszka, and M. Reimherr (2017). Testing separability of space-time functional processes. Biometrika 104(2), 425–437.

Danaher, P., P. Wang, and D. M. Witten (2014). The joint graphical lasso for inverse covariance estimation across multiple classes. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 76(2), 373–397.

Di, C.-Z., C. M. Crainiceanu, B. S. Caffo, and N. M. Punjabi (2009). Multilevel functional principal component analysis. The annals of applied statistics 3(1), 458.

Fuentes, M. (2006). Testing for separability of spatial-temporal covariance functions. Journal of statistical planning and inference 136(2), 447–466.

Happ, C. and S. Greven (2018). Multivariate functional principal component analysis for data observed on different (dimensional) domains. Journal of the American Statistical Association 113(522), 649–659.

Hsing, T. and R. Eubank (2015). Theoretical foundations of functional data analysis, with an introduction to linear operators. John Wiley & Sons.

Hu, X. and F. Yao (2022). Sparse functional principal component analysis in high dimensions. Statistica Sinica 32, 1–22.

Jiang, C.-R. and J.-L. Wang (2010). Covariate adjusted functional principal components analysis for longitudinal data. The Annals of Statistics 38(2), 1194 – 1226.

Lei, Y., Z. Zhang, and J. Jin (2010). Automatic tonnage monitoring for missing part detection in multi-operation forging processes. Journal of manufacturing science and engineering 132(5).

Li, B., M. G. Genton, and M. Sherman (2007). A nonparametric assessment of properties of space–time covariance functions. Journal of the American Statistical Association 102(478), 736–744.

Li, Y. and Y. Guan (2014). Functional principal component analysis of spatio-temporal point processes with applications in disease surveillance. Journal of the American Statistical Association 109(507), 1205–1215.

Liang, D., H. Huang, Y. Guan, and F. Yao (2022). Test of weak separability for spatially stationary

functional field. Journal of the American Statistical Association, 1–14.

Liu, C., S. Ray, and G. Hooker (2017). Functional principal component analysis of spatially correlated data. Statistics and Computing 27(6), 1639–1654.

Lynch, B. and K. Chen (2018). A test of weak separability for multi-way functional data, with application to brain connectivity studies. Biometrika 105(4), 815–831.

Masak, T., S. Sarkar, and V. Panaretos (2023). Separable expansions for covariance estimation via the partial inner product. Biometrika 110(1), 225–247.

Paynabar, K., C. Zou, and P. Qiu (2016). A change-point approach for phase-i analysis in multivariate profile monitoring and diagnosis. Technometrics 58(2), 191–204.

Qiao, X., S. Guo, and G. M. James (2019). Functional graphical models. Journal of the American Statistical Association 114(525), 211–222.

Sherman, M. (2011). Spatial Statistics and Spatio-temporal Data: Covariance Functions and Directional Properties. John Wiley & Sons.

Simpson, S. L., L. J. Edwards, M. A. Styner, and K. E. Muller (2014). Separability tests for high-dimensional, low-sample size multivariate repeated measures data. Journal of applied statistics 41(11), 2450–2461.

Wikle, C. K., A. Zammit-Mangion, and N. Cressie (2019). Spatio-temporal Statistics with R. Chapman and Hall/CRC.

Xu, G., M. Wang, J. Bian, H. Huang, T. R. Burch, S. C. Andrade, J. Zhang, and Y. Guan (2020). Semi-parametric learning of structured temporal point processes. The Journal of Machine Learning Research 21(1), 7851–7889.

Xue, K. and F. Yao (2020). Distribution and correlation-free two-sample test of high-dimensional means. The Annals of Statistics 48(3), 1304 – 1328.

Yao, F., H.-G. Müller, and J.-L. Wang (2005, June). Functional data analysis for sparse longitudinal data. Journal of the American Statistical Association 100(470), 577–590.

Zapata, J., S.-Y. Oh, and A. Petersen (2022). Partial separability and functional graphical models for multivariate gaussian processes. Biometrika 109(3), 665–681.

Zhang, H. and Y. Li (2021). Unified principal component analysis for sparse and dense functional data under spatial dependency. Journal of Business & Economic Statistics, 1–15.

Zhang, J.-T. (2013). Analysis of variance for functional data.

Zhang, X. L., H. Begleiter, B. Porjesz, W. Wang, and A. Litke (1995). Event related potentials during object recognition tasks. Brain research bulletin 38(6), 531–538.

School of Mathematics, Sun Yat-sen University, P.R.China

E-mail: luofzh3@mail2.sysu.edu.cn

Department of Probability and Statistics, School of Mathematical Sciences, Center for Statistical Science, Peking University, Beijing, P.R.China;

E-mail: 2001110073@stu.pku.edu.cn

School of Statistics and Data Science, Nankai University, P.R.China

E-mail: liangdecai@nankai.edu.cn