Journal of Multivariate Analysis 189 (2022) 104887

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Extremal dependence measure for functional data

Mihyun Kim [1], Piotr Kokoszka [1,*]

Department of Statistics, Colorado State University, Fort Collins, CO 80523-1877, United States

ARTICLE INFO

Article history:
Received 21 October 2020
Received in revised form 16 September 2021
Accepted 20 October 2021
Available online 3 November 2021

AMS 2010 subject classifications:
primary 62G05
secondary 62G32
tertiary 62H20

Keywords:
Functional data
Functional scores
Multivariate extremal dependence
Regular variation

ABSTRACT

Principal component analysis is one of the most fundamental tools of functional data analysis. It leads to an efficient representation of infinitely dimensional objects, like curves, by means of multivariate vectors of scores. We study the dependence between extremal values of the scores using the extremal dependence measure (EDM). The EDM has been proposed and studied for positive bivariate observations. After extending it to multivariate observations, we focus on its application to the vectors of scores of functional data. Estimated scores form a triangular array of dependent random variables. We derive condition guaranteeing that a suitable estimator of the EDM based on these scores converges to the population EDM and is asymptotically normal. These conditions are completely different from those encountered in the second-order theory of functional data. They are formulated within the framework of functional regular variation. Large sample theory is complemented by an application to intraday return curves for certain stocks and by a simulation study.

1. Introduction

We first concisely state main contributions of the paper with the caveat that detailed definitions and formulations will be provided in the following. Consider a sample of functions $X_i(t)$, $t \in \mathcal{T}$, such that each of them has the same distribution as X. The Karhunen–Loéve expansion is $X(t) = \sum_{j=1}^{\infty} \xi_j v_j(t)$. The functions v_j are the functional principal components (FPCs) and the random variables ξ_j are their scores. We want to estimate extremal dependence of ξ_j and $\xi_{j'}$. We define a measure of such a dependence, which we denote by $\mathrm{D}(\xi_j, \xi_{j'})$. We then define an estimator of $\mathrm{D}(\xi_j, \xi_{j'})$ and formulate conditions under which it is consistent (Theorem 1) and asymptotically normal (Theorem 2). The main difficulty is that the population scores $\xi_{ij} = \langle X_i, v_j \rangle$ are not observable.

This paper thus makes a contribution at the nexus of functional data analysis (FDA) and extreme value theory (EVT). We assume that the reader is familiar with mathematical foundations of functional data analysis and central principles of extreme value theory. The FDA background given in Chapters 2 and 3 of Horváth and Kokoszka [17] is sufficient, and more detailed treatment is provided in Hsing and Eubank [18]. Recent advances in FDA are surveyed in Goia and Vieu [15], Aneiros et al. [1], and Cuevas [5].

Chapters 2 and 6 of Resnick [31] provide sufficient background in extreme value theory. Other references are cited when needed. We assume that all functions are elements of the space $L^2 = L^2(\mathcal{T})$, where the measure space $\mathcal{T}$ is such that $L^2(\mathcal{T})$, with the usual inner product, is a separable Hilbert space. This will be ensured if the measure on $\mathcal{T}$ is σ-finite and defined on a countably generated σ-algebra, see e.g. Proposition 3.4.5 in Cohn [2]. In particular, $\mathcal{T}$ can be taken to be a complete separable metric space (Polish space).

* Corresponding author.
E-mail address: Piotr.Kokoszka@colostate.edu (P. Kokoszka).

1 Both Authors, Mihyun Kim and Piotr Kokoszka, contributed equally to this work.

https://doi.org/10.1016/j.jmva.2021.104887

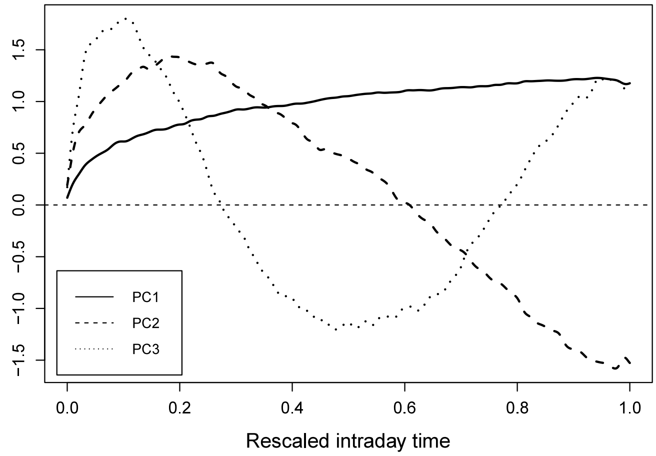

Fig. 1. The first three sample FPCs of intraday returns on Walmart stock based on sample of 1378 curves.

Suppose $X_1, \ldots, X_n$ are mean zero iid functions in L^2 with $\mathrm{E}\|X_i\|^2 < \infty$, and denote by X a generic random function with the same distribution as each X_i. A main dimension reduction tool of functional data analysis is to project the infinite dimensional functions X_i onto a finite dimensional subspace spanned by the FPCs. We now recall the required definitions. Consider the population covariance operator of X, defined by

$$C(x) := \mathrm{E}[\langle X, x\rangle X], \quad x \in L^2. \tag{1}$$

The eigenfunctions of C are the FPCs, denoted by $v_j, j \geq 1$, i.e., $C(v_j) = \lambda_j v_j$, where the λ_j are the eigenvalues of C. The FPCs lead to the commonly used Karhunen–Loéve expansion

$$X_i(t) = \sum_{j=1}^{\infty} \xi_{ij} v_j(t), \quad \xi_{ij} = \left\langle X_i, v_j \right\rangle, \quad \mathrm{E}\xi_{ij}^2 = \lambda_j. \tag{2}$$

The FPCs v_j and the eigenvalues λ_j are estimated by $\hat{v}_j$ and $\hat{\lambda}_j$, which are solutions to the equations

$$\widehat{C}(\hat{v}_j)(t) = \hat{\lambda}_j \hat{v}_j(t), \quad \text{foralmostall} t \in \mathcal{T}, \tag{3}$$

where $\widehat{C}$ is the sample covariance operator defined by

$$\widehat{C}(x)(t) = \frac{1}{n} \sum_{i=1}^{n} \langle X_i, x\rangle X_i, \quad x \in L^2.$$

Each curve X_i can then be approximated by a linear combination of a finite set of the estimated FPCs $\hat{v}_j$, i.e., $X_i(t) \approx \sum_{j=1}^{p} \hat{\xi}_{ij} \hat{v}_j(t)$, where the $\hat{\xi}_{ij} = \left\langle X_i, \hat{v}_j \right\rangle$ are the sample scores. Each $\hat{\xi}_{ij}$ quantifies the contribution of the curve $\hat{v}_j$ to the shape of the curve X_i. Thus, the vector of the sample scores, $[\hat{\xi}_{i1}, \ldots, \hat{\xi}_{ip}]^\top$, encodes the shape of X_i to a good approximation. To illustrate, Fig. 1 displays the first three sample FPCs, $\hat{v}_1, \hat{v}_2, \hat{v}_3$, for intraday return curves R_i, $1 \leq i \leq 1378$, for Walmart stock from July 05, 2006 to Dec 30, 2011. These data are described in detail in Section II of the supplement. The curves R_i show how a return on an investment changes throughout a trading day as two examples are shown in Fig. 2. The curve $\hat{v}_1$ is a monotonic trend throughout the day. If the score corresponding to it is large, trading in this stock on a given day was dominated by a systematic increase (or decline if the score is negative) in the price of the stock. Notice the gradually decreasing slope of $\hat{v}_1$, which reflects the well-known fact that the most intense trading takes place after the opening of the trading floor. The second FPC, $\hat{v}_2$, has a large score, if there is a significant reversal in investor sentiment during a given trading day. These observations are illustrated in Fig. 2.

The main interest in this paper is the estimation of extremal dependence between the scores corresponding to different FPCs. Extremal dependence is a tendency of large values of one component to be coupled with large values of another component of a random vector. In the context of our Walmart stock example, extreme dependence between the first and second scores indicates that an extremely high monotonic trend and a pronounced reversion tend to occur simultaneously. We assess extremal dependence of the scores by means of the extremal dependence measure (EDM), which is constructed based on the theory of heavy-tailed regularly-varying random vectors. There has been considerable research on quantifying the tail dependence between extreme values in a heavy-tailed random vector. Ledford and Tawn [23,24,25] defined the coefficient of tail dependence, which was later generalized to the extremogram by Davis and Mikosch [7]. While these approaches are essentially based on the exponent measure of a random vector, the EDM is

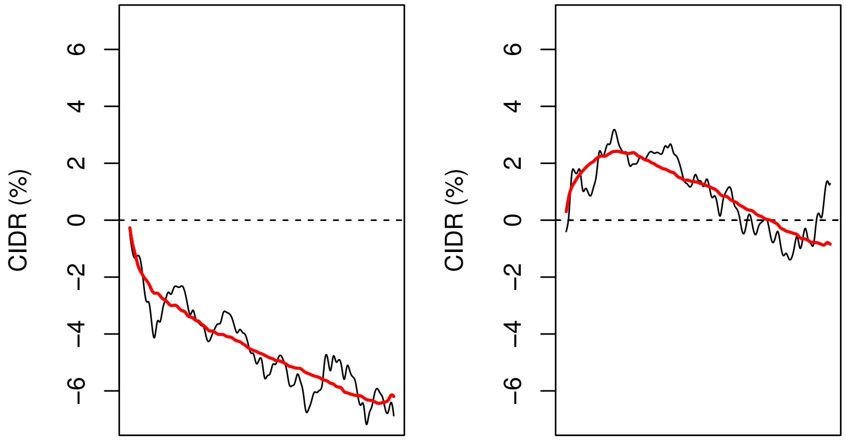

Fig. 2. Walmart intraday cumulative return curves on two trading days and their approximations by $\sum_{i=1}^{3}\hat{\xi}_{ij}\hat{v}_j(t)$. In the left panel, $\hat{\xi}_1 = -4.7, \hat{\xi}_2 = 0.4, \hat{\xi}_3 = -0.1$, observed on October 7, 2008. In the right panel, $\hat{\xi}_1 = 0.8, \hat{\xi}_2 = 1.2, \hat{\xi}_3 = 0.1$, observed on November 18, 2008.

defined in terms of the spectral measure. The EDM was introduced by Resnick [30] and further investigated by Larsson and Resnick [22]. Important related papers are Genton et al. [14] and Cooley and Thibaud [3].

In this paper, we quantify extremal dependence of scores using the EDM. To estimate the EDM of population scores, we consider an extension of the estimator proposed by Larsson and Resnick [22]. It is important to emphasize that in our functional setting, the estimator can only be computed using the sample scores $\hat{\xi}_{ij} = \langle X_i, \hat{v}_j\rangle$, not the population scores $\xi_{ij} = \langle X_i, v_j\rangle$ because the ξ_{ij} are unobservable. Establishing large sample properties of any estimator based on sample scores requires taking the effect of the estimation of the scores into account. Since the estimator $\widehat{C}$ in (3) depends on the whole sample $X_1, \ldots, X_n$, the vectors $[\hat{\xi}_{i1}, \ldots, \hat{\xi}_{ip}]^\top$ are no longer independent, even if $X_1, \ldots, X_n$ are i.i.d functions. They form a triangular array of dependent identically distributed vectors of dimension p. We also note that the population scores satisfy $\mathrm{Cov}(\xi_{ij}, \xi_{ij'}) = 0$ if $j \neq j'$ and the sample correlation of the sample scores $\hat{\xi}_{ij}$ and $\hat{\xi}_{ij'}$ is also zero. However, the correlation is a measure of the overall dependence, and there may be strong dependence, e.g. between the positive parts ξ_{ij}^+ and $\xi_{ij'}^+$, in particular there may be extremal dependence in specific quadrants. Another point to keep in mind is that for regularly varying observations, zero covariance does not imply independence.

The remainder of the paper is organized as follows. In Section 2, we introduce preliminaries on multivariate regular variation and the EDM, and extend the concept of the EDM to multivariate data. Our main large sample results are presented in Section 3, which deals with the EDM for scores of functional observations. Section 4 presents a number of preliminary results. These results allow us to streamline the exposition of the proofs of the results of Section 3, which are presented in Section 5.

The paper is accompanied by online Supplementary Material, which contains several sections. Section 1 explains how to normalize tail indexes of components of multivariate vectors. This is a well-researched topic in EVT, but may be less known in the FDA community, so a brief account needed to understand the application in Section 2 of the supplement is provided. Sections 2 and 3, present, respectively, an application to functional return data and a simulation study. Section 4 contains additional tables discussed in Section 3.

We hope that this work will be received with some interest by researchers working in two exciting and dynamic fields: functional data analysis and extreme value theory.

2. Multivariate regular variation and the EDM

We start by introducing multivariate regular variation for random vectors with positive components because the extremal dependence measure (EDM) was defined in such context. Following Resnick [31], we denote by $\mathbb{E}_d = [0, \infty]^d \setminus \{\mathbf{0}\}$ the nonnegative orthant compactified at infinity. We denote by $M_+(\mathbb{E}_d)$ the space of Radon measures on $\mathbb{E}_d$, and by $\xrightarrow{v}$ the vague convergence in $M_+(\mathbb{E}_d)$. An $\mathbb{E}_d$-valued random vector $\mathbf{Z} = [Z_1, \ldots, Z_d]^\top$ with distribution function F is regularly varying with index $-\alpha$, $\alpha > 0$, if there exist a sequence $b(n) \to \infty$ and a Radon measure ν on $\mathbb{E}_d$ such that $\nu(t\cdot) = t^{-\alpha}\nu(\cdot)$, and

$$n\,\mathrm{Pr}\left(\frac{\mathbf{Z}}{b(n)} \in \cdot\right) \xrightarrow{v} \nu, \quad \text{in } M_+(\mathbb{E}_d). \tag{4}$$

Unless stated otherwise, all limits are taken as $n \to \infty$. We assume that one-dimensional marginal distributions of ν are nondegenerate. In (4), all components are normalized by the same sequence $\{b(n)\}$, which means that all marginal distributions are tail equivalent with the index $-\alpha$, see Remark 6.1 in Resnick [31]. A possible choice for $b(n)$ is the quantile function, defined by $\Pr(Z_1 > b(n)) = n^{-1}$. When $b(n) = n$, all marginal distributions are tail equivalent to the standard Pareto distribution with $\alpha = 1$, which is called the standard case.

There are various equivalent formulations of multivariate regular variation, see Theorem 6.1 of Resnick [31]. The formulation with a polar coordinate representation is commonly used due to its computational convenience and intuitive interpretation. Fix a norm $\|\cdot\|$ in $\mathbb{R}^d$, and set $\mathbb{S}_+^d = \{x \in \mathbb{R}^d : \|x\| = 1\} \cap \mathbb{E}_d$, the unit sphere in the nonnegative orthant. A d-dimensional random vector $\mathbf{Z} = [Z_1, \ldots, Z_d]^\top$ is regularly varying if and only if there exist a sequence $b_R(n) \to \infty$ and an angular probability measure Γ on $\mathbb{S}_+^d$ such that for $(R, \Theta) = (\|\mathbf{Z}\|, \mathbf{Z}/\|\mathbf{Z}\|)$,

$$n \Pr\left(\left(\frac{R}{b_R(n)}, \Theta\right) \in \cdot\right) \xrightarrow{v} c\nu_\alpha \times \Gamma, \quad \text{in } M_+((0,\infty] \times \mathbb{S}_+^d), \tag{5}$$

where $\nu_\alpha(x,\infty] = x^{-\alpha}$ and $c = \nu\{\mathbf{x} : \|\mathbf{x}\| > 1\} > 0$. The sequence $\{b_R(n)\}$ in (5) is defined by $\Pr(R > b_R(n)) = n^{-1}$, so in this case $b_R(\cdot)$ depends on the choice of the norm $\|\cdot\|$. Definitions (4) and (5) can be extended directly to an $\mathbb{R}^d$-valued random vector with ν on $\mathbb{R}^d \setminus \{\mathbf{0}\}$ and Γ on $\mathbb{S}^d = \{x \in \mathbb{R}^d : \|x\| = 1\}$, see, e.g., Propositions 2.2.5 and 2.2.6 of Meiguet [27]. In practice, the components of a random vector might not be tail equivalent. The case of different tail indexes of the coordinates, and transformations which make the coordinates tail equivalent are discussed in Section I of the supplement.

We now turn to the EDM. Given a regularly varying nonnegative bivariate random vector $\mathbf{Z} = [Z_1, Z_2]^\top$, Larsson and Resnick [22] define the EDM by

$$\mathrm{EDM}(Z_1, Z_2) = \int_{\mathbb{S}_+^2} a_1 a_2 \Gamma(d\mathbf{a}). \tag{6}$$

The EDM takes the minimal value of zero, $\mathrm{EDM}(Z_1, Z_2) = 0$, iff the coordinates of $\mathbf{Z}$ are asymptotically independent. This means that the angular measure Γ concentrates on $\{(1,0)/\|(1,0)\|, (0,1)/\|(0,1)\|\}$, or equivalently, the exponent measure ν concentrates on the axes. Also, if the norm is symmetric, $\mathrm{EDM}(Z_1, Z_2)$ achieves its maximal value iff the distribution of $\mathbf{Z}$ has asymptotic full dependence; i.e., Γ has mass on $\{(1,1)/\|(1,1)\|\}$, or equivalently, ν concentrates on the line $\{t(1,1), t > 0\}$.

Larsson and Resnick [22] show that the EDM can be interpreted as the limit of cross moments between normalized Z_1 and Z_2 conditional on large values of $R = \|\mathbf{Z}\|$;

$$\mathrm{EDM}(Z_1, Z_2) = \lim_{r\to\infty} \mathrm{E}\left[\frac{Z_1}{R}\frac{Z_2}{R}\middle| R > r\right].$$

Based on this relation, they propose an estimator for $\mathrm{EDM}(Z_1, Z_2)$, defined by

$$\mathrm{D}_n(Z_1, Z_2) = \frac{1}{k}\sum_{i=1}^{n} \frac{Z_{i1}}{R_i}\frac{Z_{i2}}{R_i} I_{R_i \ge R_{(k)}}, \tag{7}$$

where $\mathbf{Z}_i = [Z_{i1}, Z_{i2}]^\top$, $1 \le i \le n$ are iid copies of $\mathbf{Z} = [Z_1, Z_2]^\top$, $R_i = \|\mathbf{Z}_i\|$, and $R_{(k)}$ is the kth largest order statistics with the convention $R_{(1)} = \max\{R_1, \ldots, R_n\}$.

Larsson and Resnick [22] consider non-negative bivariate vectors. To be able to work with the vectors of scores of functional data, we first have to extend their definitions to a setting of multivariate random vectors of an arbitrary dimension. Our first objective is to generalize (6) to a d-dimensional vector $\mathbf{Z} = [Z_1, \ldots, Z_d]^\top$. We formulate the EDM between the components Z_1 and Z_2 for simplicity. We first assume that all components are positive. Given the angular measure Γ on $\mathbb{S}_+^d$ for $\mathbf{Z}$, we define the EDM for Z_1 and Z_2 as

$$\mathrm{D}(Z_1, Z_2) = \int_{\mathbb{S}_+^d} \frac{a_1 a_2}{\|(a_1, a_2, 0, \ldots, 0)\|^2} \Gamma(d\mathbf{a}). \tag{8}$$

We set $a_1 a_2 / \|(a_1, a_2, 0, \ldots, 0)\|^2 = 0$ when $a_1 = a_2 = 0$. Definition (8) is different from a simple extension of (6) given by

$$\mathrm{D}'(Z_1, Z_2) = \int_{\mathbb{S}_+^d} a_1 a_2 \Gamma(d\mathbf{a}). \tag{9}$$

We will now argue that for a d-dimensional vector $\mathbf{Z}$, with $d \ge 3$, D is a better measure for assessing extremal dependence between Z_1 and Z_2 than D′. Suppose that a random vector $\mathbf{Z} = [Z_1, Z_2, Z_3]^\top$ is regularly varying with an angular measure Γ on $\mathbb{S}_+^3$, and fix the Euclidean norm $\|\cdot\|$ in $\mathbb{R}_+^3$. Consider the following four cases:

1. The angular measure Γ_1 has unit mass on $(1,1,10)/\sqrt{102}$; the exponent measure ν_1 concentrates on $\{t(1,1,10), t > 0\}$.
2. The angular measure Γ_2 has unit mass on $(1,1,1)/\sqrt{3}$; the exponent measure ν_2 concentrates on $\{t(1,1,1), t > 0\}$.
3. The angular measure Γ_3 has unit mass on $(7,7,2)/\sqrt{102}$; the exponent measure ν_3 concentrates on $\{t(7,7,2), t > 0\}$.
4. The angular measure Γ_4 has mass 1/2 on each $(1,1,10)/\sqrt{102}$ and $(7,7,2)/\sqrt{102}$; the exponent measure ν_4 concentrates on $\{t(1,1,10), t > 0\} \cup \{t(7,7,2), t > 0\}$.

Suppose Z has a Pareto distribution with index $\alpha > 0$. The following random vectors have extremal distribution corresponding to each of the above cases:

$$\mathbf{Z}^{(1)} = [Z,\ Z,\ 10Z], \quad \mathbf{Z}^{(2)} = [Z,\ Z,\ Z], \quad \mathbf{Z}^{(3)} = [7Z,\ 7Z,\ 2Z],$$

$$\mathbf{Z}^{(4)} = \xi[Z,\ Z,\ 10Z] + (1-\xi)[7Z,\ 7Z,\ 2Z],$$

where ξ is a Bernoulli random variable with probability of success $1/2$.

Set $\mathbb{P}_{12} = \{[t_1, t_2, 0], t_1, t_2 \in \mathbb{R}\}$. The projections of the random vectors $\mathbf{Z}^{(1)}, \mathbf{Z}^{(2)}, \mathbf{Z}^{(3)}$, and $\mathbf{Z}^{(4)}$ onto $\mathbb{P}_{12}$ are, respectively,

$$\widetilde{\mathbf{Z}}^{(1)} = [Z,\ Z], \quad \widetilde{\mathbf{Z}}^{(2)} = [Z,\ Z], \quad \widetilde{\mathbf{Z}}^{(3)} = [7Z,\ 7Z], \quad \widetilde{\mathbf{Z}}^{(4)} = [\xi Z + 7(1-\xi)Z,\ \xi Z + 7(1-\xi)Z].$$

For all of the projected random vectors, the two components are equal, so a good measure of extremal dependence between them should attain its maximal value. Since we use the Euclidean norm and Γ is normalized to unity, the maximum value of both D and D′ is 1/2. Direct verification shows that we achieve the maximum value for all cases using the measure D. The measure D′ however does not give the maximum value. For each case:

$$\mathrm{D}'(Z_1^{(1)}, Z_2^{(1)}) = \frac{1}{102}, \quad \mathrm{D}'(Z_1^{(2)}, Z_2^{(2)}) = \frac{34}{102},$$
$$\mathrm{D}'(Z_1^{(3)}, Z_2^{(3)}) = \frac{49}{102}, \quad \mathrm{D}'(Z_1^{(4)}, Z_2^{(4)}) = \frac{1}{102}\frac{1}{2} + \frac{49}{102}\frac{1}{2} = \frac{25}{102}.$$

It can be further shown that, for any norm $\|\cdot\|$ in $\mathbb{R}^d$, the measures D and D′, defined for d-dimensional vector $\mathbf{Z}$ with $d \geq 3$, are not equivalent in the sense of Definition 1 on p.234 of Larsson and Resnick [22], which we now recall. For a given $\mathbf{Z}$, let $\rho_i(\mathbf{Z}) = \int_{\mathbb{S}_+^d} k_i(\mathbf{a})\Gamma(d\mathbf{a})$ for some nonnegative map $k_i : \mathbb{S}_+^d \mapsto \mathbb{R}_+$. Then $\rho_1(\mathbf{Z})$ and $\rho_2(\mathbf{Z})$ are equivalent if and only if there are constants $0 < m \leq M < \infty$ such that

$$m\rho_1(\mathbf{Z}) \leq \rho_2(\mathbf{Z}) \leq M\rho_1(\mathbf{Z}).$$

It is obvious that the measures D and D′ are equivalent for a bivariate vector $\mathbf{Z}$. We formalize the nonequivalence between the measures for a d-dimensional vector $\mathbf{Z}$ with $d \geq 3$ in the following proposition.

Proposition 1. *Suppose that a $\mathbb{E}_d$-valued random vector $\mathbf{Z} = [Z_1, \ldots, Z_d]^\top$ is regularly varying with angular measure Γ on $\mathbb{S}_+^d$, with $d \geq 3$. Then $\mathrm{D}(Z_1, Z_2)$ and $\mathrm{D}'(Z_1, Z_2)$, defined in (8), (9), respectively, are not equivalent for any norm $\|\cdot\|$ in $\mathbb{R}^d$.*

Proof. Proposition 1 of Larsson and Resnick [22] shows that $\rho_1(\mathbf{Z})$ and $\rho_2(\mathbf{Z})$ are equivalent if and only if there are constants $0 < m \leq M < \infty$ such that

$$mk_1(\mathbf{a}) \leq k_2(\mathbf{a}) \leq Mk_1(\mathbf{a}), \quad \forall\, \mathbf{a} \in \mathbb{S}_+^d. \tag{10}$$

Observe that the ratio of the integrand in $\mathrm{D}'(Z_1, Z_2)$ to the integrand in $\mathrm{D}(Z_1, Z_2)$ is $\|(a_1, a_2, 0, \ldots, 0)\|^2$. This ratio is clearly zero at $\mathbf{a} = \mathbf{0}$, violating (10), but $\mathbf{0} \notin \mathbb{S}_+^d$. We therefore consider a path in $\mathbb{S}_+^d$ defined by

$$\mathbf{a}(x) = (x, x, 1, 0, \ldots, 0)/\|(x, x, 1, 0, \ldots, 0)\|, \quad x \searrow 0.$$

Then,

$$\|(a_1(x), a_2(x), 0, 0, \ldots, 0)\|^2 = \frac{\|(x, x, 0, 0, \ldots, 0)\|^2}{\|(x, x, 1, 0, \ldots, 0)\|^2} \to 0,$$

as $x \searrow 0$ because every norm in $\mathbb{R}^d$ is equivalent to the Euclidean norm. □

Another question of interest is the relationship between $\mathrm{D}(Z_1, Z_2)$ in (8) and $\mathrm{EDM}(Z_1, Z_2)$ in (6). We clarify it in the following proposition. Related results are derived in Opitz [28], de Fondeville [12] and de Fondeville and Davison [13].

Proposition 2. *Suppose that the exponent measure and angular measure of a d-dimensional regularly-varying random vector $\mathbf{Z} = [Z_1, \ldots, Z_d]^\top$ are, respectively, ν on $\mathbb{E}_d$ and Γ on $\mathbb{S}_+^d$. Denote the exponent measure and angular measure of the bivariate vector $[Z_1, Z_2]^\top$, respectively, by ν_2 on $\mathbb{E}_2$ and Γ_2 on $\mathbb{S}_+^2$. Then,*

$$\mathrm{D}(Z_1, Z_2) = \int_{\mathbb{S}_+^d} \frac{a_1 a_2}{\|(a_1, a_2, 0, \ldots, 0)\|^2}\Gamma(d\mathbf{a}) = \int_{\mathbb{S}_+^2} b_1 b_2 \Gamma_2(d\mathbf{b}) = \mathrm{EDM}(Z_1, Z_2)$$

and, for any Borel set $G \subset \mathbb{E}_2$,

$$\nu_2(G) = \nu(G \times [0, \infty]^{d-2}).$$

Proof. We first clarify the connection between the measure ν on $\mathbb{E}_d$ and the measure ν_2 on $\mathbb{E}_2$. By (4), for any measurable rectangle $A \times B \subset \mathbb{E}_2$,

$$\frac{\nu(A \times B \times [0,\infty]^{d-2})}{\nu_2(A \times B)} = \lim_{n\to\infty} \frac{n \Pr\left(\mathbf{Z}/b(n) \in A \times B \times [0,\infty]^{d-2}\right)}{n \Pr\left(Z_1/b(n) \in A, Z_2/b(n) \in B\right)} = 1.$$

We conclude that the measure ν_2 is obtained by integrating the entire measure ν over all coordinates except for the first two.

According to formulas on page 239 of Larsson and Resnick [22], $\mathrm{EDM}(Z_1, Z_2)$ can be expressed as

$$\int_{\mathbb{S}^2_+} b_1 b_2 \Gamma_2(d\mathbf{b}) = \frac{1}{\nu_2\left(\|(y_1,y_2)\| > 1\right)} \int_{\|(y_1,y_2)\|>1} \frac{y_1 y_2}{\|(y_1,y_2)\|^2}\, \nu_2(dy_1 dy_2).$$

Therefore, using the relationship between ν_2 and ν,

$$\int_{\mathbb{S}^2_+} b_1 b_2 \Gamma_2(d\mathbf{b}) = \frac{1}{\nu\left(\{\mathbf{y} : \|(y_1,y_2,0,\dots,0)\| > 1\}\right)} \int_{\{\mathbf{y}:\|(y_1,y_2,0,\dots,0)\|>1\}} \frac{y_1 y_2}{\|(y_1,y_2,0,\dots,0)\|^2}\, \nu(d\mathbf{y}).$$

Applying the polar transformation T defined by $T(\mathbf{y}) = (\|\mathbf{y}\|, \mathbf{y}/\|\mathbf{y}\|)$ for $\mathbf{y} \in \mathbb{E}_d$, we obtain

$$\int_{\mathbb{S}^2_+} b_1 b_2 \Gamma_2(d\mathbf{b}) = \frac{1}{\nu\left(\{\mathbf{y} : \|(y_1,y_2,0,\dots,0)\| > 1\}\right)} \int_{T(\{\mathbf{y}:\|(y_1,y_2,0,\dots,0)\|>1\})} f \circ T^{-1}(r,\mathbf{a})\, \nu \circ T^{-1}(dr \times d\mathbf{a}),$$

where $f(\mathbf{y}) = y_1 y_2/\|(y_1,y_2,0,\dots,0)\|^2$. First observe that

$$\begin{aligned} T\left(\{\mathbf{y} : \|(y_1,y_2,0,\dots,0)\| > 1\}\right) &= \{(r,(a_1,a_2,\dots,a_d)) : \|(ra_1, ra_2, 0,\dots,0)\| > 1\} \\ &= \{(r,(a_1,a_2,\dots,a_d)) : r > \|(a_1,a_2,0,\dots,0)\|^{-1}\}. \end{aligned}$$

Using the fact that $\nu \circ T^{-1} = c\nu_\alpha \times \Gamma$, where $c = \nu(\|\mathbf{y}\| > 1)$, we obtain

$$\begin{aligned} &\nu\left(\{\mathbf{y} : \|(y_1,y_2,0,\dots,0)\| > 1\}\right) = \nu \circ T^{-1}\left(T\left(\{\mathbf{y} : \|(y_1,y_2,0,\dots,0)\| > 1\}\right)\right) \\ &= c\nu_\alpha \times \Gamma\left(\{(r,(a_1,a_2,\dots,a_d)) : r > \|(a_1,a_2,0,\dots,0)\|^{-1}\}\right) = c\,\|(a_1,a_2,0,\dots,0)\|^\alpha. \end{aligned}$$

Therefore,

$$\begin{aligned} \int_{\mathbb{S}^2_+} b_1 b_2 \Gamma_2(d\mathbf{b}) &= \frac{1}{c\,\|(a_1,a_2,0,\dots,0)\|^\alpha} \int_{\mathbb{S}^d_+} \int_{r>\|(a_1,a_2,0,\dots,0)\|^{-1}} \frac{a_1 a_2}{\|(a_1,a_2,0,\dots,0)\|^2}\, c\nu_\alpha(dr)\Gamma(d\mathbf{a}) \\ &= \int_{\mathbb{S}^d_+} \frac{a_1 a_2}{\|(a_1,a_2,0,\dots,0)\|^2}\, \Gamma(d\mathbf{a}). \quad \square \end{aligned}$$

By Proposition 2 we can use the estimator (7), originally introduced for $\mathrm{EDM}(Z_1, Z_2)$, to estimate $\mathrm{D}(Z_1, Z_2)$ as well.

A further extension of the EDM (6) is that from the nonnegative quadrant to the four quadrants, as a vector of the scores takes on values in $\mathbb{R}^d$. Larsson and Resnick [22] define the EDM for a nonnegative random vector, but (6) can be readily generalized to a random vector $\mathbf{Z} = [Z_1, Z_2]^\top$ with real components. Suppose that $\mathbf{Z} = [Z_1, Z_2]^\top$ in $\mathbb{R}^2$ is regularly varying with an angular measure Γ_2 on $\mathbb{S}^2$. Then, we define the EDM for $\mathbf{Z} = [Z_1, Z_2]^\top$ by

$$\mathrm{EDM}(Z_1, Z_2) = \int_{\mathbb{S}^2} a_1 a_2 \Gamma_2(d\mathbf{a}). \tag{11}$$

The above definition allows us to quantify the strength of the extremal dependence between Z_1 and Z_2 in $\mathbb{R}^2$. Unlike (6), (11) can take a negative value depending on which quadrants Γ_2 has its mass on, so careful interpretation is needed. To explore the dependence spectrum that (11) can measure, we fix the Euclidean norm $\|\cdot\|$ in $\mathbb{R}^2$. Then, (11) has a range from $-1/2$ to $1/2$. The maximal value, $1/2$, indicates a perfect positive extremal dependence; here, "positive" means that Z_1 and Z_2 have the same signs, and "perfect" means that the magnitudes of Z_1 and Z_2 show asymptotic full dependence, i.e., Γ_2 concentrates on $\{(1,1)/\sqrt{2},\ (-1,-1)/\sqrt{2}\}$. Similarly, the minimum value, $-1/2$, indicates a perfect negative extremal dependence; "negative" means that Z_1 and Z_2 have the opposite signs, and in this case Γ has mass on $\{(-1,1)/\sqrt{2},\ (1,-1)/\sqrt{2}\}$.

Note that if $\mathbf{Z}$ exhibits asymptotic independence, i.e., its exponent measure concentrates on the standard axes, then (11) is 0, but the reverse does not necessarily hold true. For example, if Γ_2 concentrates equally on each element of

$$\{(1,1)/\sqrt{2},\ (-1,1)/\sqrt{2},\ (-1,-1)/\sqrt{2},\ (1,-1)/\sqrt{2}\},$$

then (11) is 0, but each quadrant shows the perfect dependence. To avoid this issue and take into account the extremal dependence in each quadrant, we suggest to complement (11) on the unit sphere $\mathbb{S}^2$ with its decomposition into the four quadrants. Let $\mathbb{S}^2_{(+,+)} = \mathbb{S}^2 \cap \{(x_1,x_2) \in \mathbb{R}^2 : x_1 \geq 0, x_2 \geq 0\}$. Similarly, let $\mathbb{S}^2_{(-,+)} = \mathbb{S}^2 \cap \{x_1 \leq 0, x_2 \geq 0\}$,

$\mathbb{S}^2_{(-,-)} = \mathbb{S}^2 \cap \{x_1 \le 0, x_2 \le 0\}$, and $\mathbb{S}^2_{(+,-)} = \mathbb{S}^2 \cap \{x_1 \ge 0, x_2 \le 0\}$. We define the supplementary measure for (11) by splitting the EDM into the four quadrant spheres,

$$\left[\int_{\mathbb{S}^2_{(+,+)}} a_1a_2\Gamma_2(d\mathbf{a}),\ \int_{\mathbb{S}^2_{(-,+)}} a_1a_2\Gamma_2(d\mathbf{a}),\ \int_{\mathbb{S}^2_{(-,-)}} a_1a_2\Gamma_2(d\mathbf{a}),\ \int_{\mathbb{S}^2_{(+,-)}} a_1a_2\Gamma_2(d\mathbf{a})\right]. \tag{12}$$

To estimate each of the components in (12), we slightly modify (7); for example, an estimator for $\int_{\mathbb{S}^2_{(+,+)}} a_1a_2\Gamma(d\mathbf{a})$ is

$$\mathrm{D}_n^{(+,+)}(Z_1,Z_2) = \frac{1}{k}\sum_{i=1}^n \frac{Z_{i1}}{R_i}\frac{Z_{i2}}{R_i} I_{R_i\ge R_{(k)}} I_{Z_{i1}\ge 0, Z_{i2}\ge 0}.$$

To elaborate, we first order the n bivariate vectors by norm and consider the top k vectors with large norm. We then use only those for which $Z_{i1} \ge 0$ and $Z_{i2} \ge 0$ from the k vectors. Estimators for the other components in (12) can be obtained in the same manner reflecting the different quadrants.

We conclude this section with an analog of Proposition 2. Given an $\mathbb{R}^d$-valued random vector $[Z_1,\dots,Z_d]^\top$, we can measure extremal dependence between Z_1 and Z_2 using (8), but integrated over the whole sphere $\mathbb{S}^d$. Following the steps in the proof of Proposition 2, it is readily shown that $\mathrm{D}(Z_1,Z_2)$ for two components of an $\mathbb{R}^d$-valued vector is in fact the same as (11).

Corollary 1. *Suppose the angular measure of a $\mathbb{R}^d$-valued random vector $[Z_1,\dots,Z_d]^\top$ is Γ on $\mathbb{S}^d$ and the angular measure of $[Z_1,Z_2]^\top$ is Γ_2 on $\mathbb{S}^2$. Then,*

$$\mathrm{D}(Z_1,Z_2) = \int_{\mathbb{S}^d} \frac{a_1a_2}{\|(a_1,a_2,0,\dots,0)\|^2}\Gamma(d\mathbf{a}) = \int_{\mathbb{S}^2} b_1b_2\Gamma_2(d\mathbf{b}) = \mathrm{EDM}(Z_1,Z_2).$$

3. The EDM for scores of functional data

In this section, we consider the estimation of the EDM of scores of functional data. Following the framework introduced in Section 1, recall that $X_1,\dots,X_n$ are mean zero iid functions in L^2 with $\mathrm{E}\,\|X_i\|^2 < \infty$, and that each X_i admits the Karhunen–Loéve expansion (2). The unknown population scores $\xi_{ij} = \langle X_i, v_j\rangle$ in (2) are estimated by the sample scores $\hat{\xi}_{ij} = \langle X_i, \hat{v}_j\rangle$, where the $\hat{v}_j$ are estimators of the FPCs v_j. We introduce the following random variables:

$$\mathbf{Y}^d = [\xi_1,\dots,\xi_d]^\top,\quad \xi_j = \langle X, v_j\rangle,\quad \mathbf{Y}_i^d = [\xi_{i1},\dots,\xi_{id}]^\top,\quad \xi_{ij} = \langle X_i, v_j\rangle,$$

$$\widehat{\mathbf{Y}}^d = [\hat{\xi}_1,\dots,\hat{\xi}_d]^\top,\quad \hat{\xi}_j = \langle X, \hat{v}_j\rangle,\quad \widehat{\mathbf{Y}}_i^d = [\hat{\xi}_{i1},\dots,\hat{\xi}_{id}]^\top,\quad \hat{\xi}_{ij} = \langle X_i, \hat{v}_j\rangle.$$

To quantify the extremal dependence between components ξ_j and $\xi_{j'}$ in $\mathbf{Y}^d$, we consider the EDM, $\mathrm{D}(\xi_j,\xi_{j'})$, defined in (8). Then, by Corollary 1,

$$\mathrm{D}(\xi_j,\xi_{j'}) = \int_{\mathbb{S}^2} a_1a_2\Gamma_{jj'}(d\mathbf{a}), \tag{13}$$

where $\Gamma_{jj'}$ on $\mathbb{S}^2$ is the angular measure of the bivariate random vector $[\xi_j,\xi_{j'}]^\top$.

Set $\mathbf{Y}_i = [\xi_{ij},\xi_{ij'}]^\top$, $\widehat{\mathbf{Y}}_i = [\hat{\xi}_{ij},\hat{\xi}_{ij'}]^\top$, $1 \le i \le n$, where we suppress the dependence of the bivariate vectors on j and j'. In light of (7), we consider two random variables that approximate $\mathrm{D}(\xi_j,\xi_{j'})$:

$$\mathrm{D}_n(\xi_j,\xi_{j'}) := \frac{1}{k}\sum_{i=1}^n \frac{\xi_{ij}}{R_i}\frac{\xi_{ij'}}{R_i} I_{R_i\ge R_{(k)}},$$

$$\widehat{\mathrm{D}}_n(\xi_j,\xi_{j'}) := \frac{1}{\hat{k}}\sum_{i=1}^n \frac{\hat{\xi}_{ij}}{\widehat{R}_i}\frac{\hat{\xi}_{ij'}}{\widehat{R}_i} I_{\widehat{R}_i\ge \widehat{R}_{(\hat{k})}}, \tag{14}$$

where $R_i = \|\mathbf{Y}_i\|$, $\widehat{R}_i = \|\widehat{\mathbf{Y}}_i\|$, and $R_{(k)}$ and $\widehat{R}_{(\hat{k})}$ are the respective largest order statistics. There is a fundamental difference between $\mathrm{D}_n(\xi_j,\xi_{j'})$ and $\widehat{\mathrm{D}}_n(\xi_j,\xi_{j'})$; $\mathrm{D}_n(\xi_j,\xi_{j'})$ is an infeasible estimator because the FPCs v_j are not observable, so the ξ_{ij} cannot be computed from the data. The estimator based on the sample scores, $\widehat{\mathrm{D}}_n(\xi_j,\xi_{j'})$, is what we can actually compute. Therefore, the consistency of $\widehat{\mathrm{D}}_n(\xi_j,\xi_{j'})$ for $\mathrm{D}(\xi_j,\xi_{j'})$ must be established. As noted in the Introduction, the sample scores $\hat{\xi}_{ij}$ are no longer independent in i (nor in j); they form a triangular array of dependent identically distributed vectors of dimension d. This new aspect of EDM estimation is specific to functional data. To handle it rigorously, we must introduce a suitable framework for regular variation of functional data. We follow Hult and Lindskog [19] and Meiguet [27].

Hult and Lindskog [19] introduced a framework based on M_0 convergence, where M_0 is the space of measures on a complete separable metric space. Meiguet [27] further investigated regular variation in Banach spaces using the notion of M_0 convergence. We define a regularly varying function in a separable Banach space $\mathbb{B}$ as follows.

Definition 1. Denote the norm in $\mathbb{B}$ by $\|\cdot\|_{\mathbb{B}}$ and the unit sphere in $\mathbb{B}$ by $\mathbb{S} := \{x \in \mathbb{B} : \|x\|_{\mathbb{B}} = 1\}$. A random element X in $\mathbb{B}$ is regularly varying with index $-\alpha, \alpha > 0$ if any of the following conditions hold:

(i) There exist a measure ν and a regularly varying sequence $b(n) \to \infty$ with index $1/\alpha$ such that

$$n \Pr\left(\frac{X}{b(n)} \in \cdot\right) \xrightarrow{M_0} \nu(\cdot), \quad n \to \infty, \tag{15}$$

where ν is a non-null measure (exponent measure) on the Borel σ-field $\mathcal{B}(\mathbb{B}_0)$ of $\mathbb{B}_0 = \mathbb{B} \setminus \{\mathbf{0}\}$.

(ii) There exist a probability measure Γ on $\mathbb{S}$ and a regularly varying sequence $b_R(n) \to \infty$ such that, for any $y > 0$,

$$n \Pr\left(\|X\|_{\mathbb{B}} > y b_R(n), X/\|X\|_{\mathbb{B}} \in \cdot\right) \xrightarrow{w} c y^{-\alpha} \Gamma(\cdot), \quad n \to \infty, \tag{16}$$

There are several equivalent definitions, see Section 2.2 of Meiguet [27], which also contains all details. The quantile function $b(t)$ in (15) admits the representation

$$b(t) = t^{1/\alpha} L(t), \quad t > 0, \tag{17}$$

where L is slowly varying as $t \to \infty$, see e.g. Resnick [31] p. 20, for the definition. An analogous representation holds for the function b_R. With the choice of $b_R(n)$, defined by $\Pr(\|X\|_{\mathbb{B}} > b_R(n)) = n^{-1}$, we get $c = 1$ in (16) since $\Gamma(\mathbb{S}) = 1$ for any $y > 0$.

We briefly review the theory of M_0 convergence. Let $B_\varepsilon := \{z \in \mathbb{B} : \|z\|_{\mathbb{B}} < \varepsilon\}$ be the open ball of radius $\varepsilon > 0$ centered at the origin. A Borel measure ν defined on $\mathbb{B}_0$ is said to be boundedly finite if $\nu(A) < \infty$, for all Borel sets that are bounded away from $\mathbf{0}$, i.e., $A \cap B_\varepsilon = \emptyset$, for some $\varepsilon > 0$. Let $\mathbb{M}_0$ be the collection of all such measures. For $\nu_n, \nu \in \mathbb{M}_0$, the ν_n converge to ν in the M_0 topology, if $\nu_n(A) \to \nu(A)$, for all bounded away from $\mathbf{0}$, ν-continuity Borel sets A, i.e., $\nu(\partial A) = 0$, where ∂A is the boundary of A. If $\mathbb{B}$ is an Euclidean space, Definition 1 is equivalent to regular variation as defined in Section 2.

We work in the Hilbert space L^2, so in the following we replace the general Banach space $\mathbb{B}$ with a separable Hilbert space $\mathbb{H}$. We define the finite-dimensional projection of $z \in \mathbb{H}$ on the subspace spanned by $f_1, \ldots, f_d \in \mathbb{H}$ by

$$\pi_{f_1,\ldots,f_d}(z) := [\langle z, f_1\rangle, \ldots, \langle z, f_d\rangle]^\top.$$

We claim in the following proposition that regular variation in $\mathbb{H}$ implies regular variation of the finite-dimensional projections in $\mathbb{R}^d$. To lighten the notation, we suppress the subscript $f_1, \ldots, f_d$ so that $\pi(z) = \pi_{f_1,\ldots,f_d}(z)$. Let $\mathcal{B}(\mathbb{S}^d)$ be the Borel σ-field on $\mathbb{S}^d$. For any set S in $\mathcal{B}(\mathbb{S}^d)$, define a set of elements in $\mathbb{H}$ by

$$\mathcal{A}_\pi(S) := \{z \in \mathbb{H} : \|\pi(z)\| > 1, \ \pi(z)/\|\pi(z)\| \in S\}. \tag{18}$$

Proposition 3. *If a random element X in $\mathbb{H}$ is regularly varying with index $-\alpha, \alpha > 0$, and $\nu(\mathcal{A}_\pi(\mathbb{S}^d)) > 0$, then $\pi(X)$ is regularly varying in $\mathbb{R}^d$ with index $-\alpha$.*

In our FDA context, the functions $f_1, \ldots, f_d$ of interest are the FPCs $v_1, \ldots, v_d$. We work under the following assumption.

Assumption 1. The functions $X_1, \ldots, X_n$ are i.i.d copies of X, which is regularly varying in L^2 according to Definition 1 with $\alpha > 2, \alpha \neq 4$. The FPCs $v_1, \ldots, v_d$ satisfy $\nu(\mathcal{A}_{\pi_{v_1,\ldots,v_d}}(\mathbb{S}^d)) > 0$ (the set $\mathcal{A}_{\pi_{v_1,\ldots,v_d}}$ is defined according to (18)).

By Proposition 3, under Assumption 1, the projection $\mathbf{Y}^d = \pi_{v_1,\ldots,v_d}(X)$ is regularly varying in $\mathbb{R}^d$ with the same index as X. The assumption $\alpha > 2$ ensures that $\mathrm{E}\|X\|^2 < \infty$, so that the FPCs can be defined. If $\alpha = 2$, then either $\mathrm{E}\|X\|^2 = \infty$ or $\mathrm{E}\|X\|^2 < \infty$ are possible, and complex assumptions on the slowly varying function L would be needed to ensure that $\mathrm{E}\|X\|^2 < \infty$. Similarly, if $\alpha = 4$, then either $\mathrm{E}\|X\|^4 = \infty$ or $\mathrm{E}\|X\|^4 < \infty$ are possible. There is a phase transition at $\alpha = 4$ found in the functional context by Kokoszka et al. [21]. The phase transitions at $\alpha = 2$ and $\alpha = 4$ in various context related to regular variation have been well-known since the 1980s, see, e.g., Theorem 3.5 in Davis and Mikosch [6], earlier papers of Davis and Resnick [8,9,10], and Embrechts et al. [11] for a broad picture. We therefore exclude $\alpha = 2$ and $\alpha = 4$ from our analysis. In the context of research on regularly varying and heavy-tailed random elements, the chief restriction is $\alpha > 2$, needed to ensure that the FPC are readily defined. It is conceivable that in the context of functions whose projections are heavy-tailed, data-driven bases different from the FPC might be appropriate, but such bases have not been devised yet.

As noted earlier, the sample scores $\hat{\xi}_{ij} = \langle X_i, \hat{v}_j\rangle$ form a triangular array whose elements are dependent across i and j. We now review bounds on the distance $\hat{v}_j - v_j$. As noted in the Introduction, these bounds apply to $\mathrm{sign}\left(\langle \hat{v}_j, v_j\rangle\right) \hat{v}_j - v_j$, but the sign always cancels in final formulas, so we assume that $\mathrm{sign}\left(\langle \hat{v}_j, v_j\rangle\right) = 1$. Recall that v_j is the jth eigenfunction of the covariance operator C in (1) corresponding to the eigenvalue λ_j, and $\hat{v}_j$ is the jth eigenfunction of its estimator $\widehat{C}$ in (3). By Lemma 2.3 in Horváth and Kokoszka [17],

$$\|\hat{v}_j - v_j\| \le A_j \|\widehat{C} - C\|_{\mathcal{L}}, \tag{19}$$

provided $d_j > 0$, where $A_j = 2\sqrt{2}/d_j$, and

$$d_1 = \lambda_1 - \lambda_2, \quad d_j = \min\left\{\lambda_{j-1} - \lambda_j, \lambda_j - \lambda_{j+1}\right\}, \ j \ge 2. \tag{20}$$

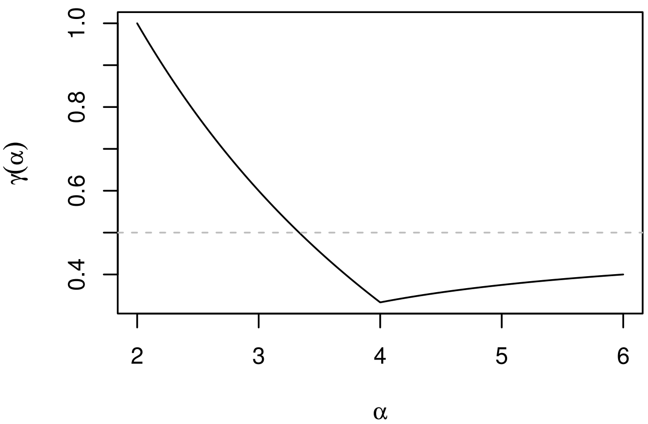

Fig. 3. The graph of the function $\gamma(\alpha)$ for $\alpha \in (2, 6)$.

The asymptotic properties of the distance between $\widehat{C}$ and C are separated into two cases depending on the range of α. If $\alpha > 4$, then $\mathrm{E}\|X\|^4 < \infty$, so, by Theorem 2.5 in Horváth and Kokoszka [17],

$$\mathrm{E}\|\widehat{C} - C\|^2 = O(n^{-1}). \tag{21}$$

Using (19), we have

$$\mathrm{E}\|\hat{v}_j - v_j\|^2 = O(n^{-1}). \tag{22}$$

The case of regularly varying X with tail index $\alpha \in (2, 4)$, which implies $\mathrm{E}\|X\|^2 < \infty$ and $\mathrm{E}\|X\|^4 = \infty$, is studied in Kokoszka et al. [21]. Under week conditions, relation (21) must be replaced by

$$\mathrm{E}\|\widehat{C} - C\|_{\mathcal{L}}^{\beta} \le L_\beta(n) n^{-\beta(1-2/\alpha)}, \quad \forall\, \beta \in (0, \alpha/2), \tag{23}$$

where L_β is a slowly varying function. For a fixed α, the strongest bound is obtained as $\beta \nearrow \alpha/2$, in which case $\beta(1 - 2/\alpha) \nearrow \alpha/2 - 1$. As $\alpha \nearrow 4$ and $\beta \nearrow \alpha/2$, relation (23) thus approaches, in a heuristic sense, relation (21). From (19) and (23), we get the condition

$$\mathrm{E}\|\hat{v}_j - v_j\|^\beta = o\left(n^{-\kappa}\right), \quad \forall\, \beta \in \left(1, \frac{\alpha}{2}\right), \ \forall\, \kappa \in \left(0, \beta\left(1 - \frac{2}{\alpha}\right)\right). \tag{24}$$

To see this, observe that

$$n^\kappa \mathrm{E}\|\hat{v}_j - v_j\|^\beta \le A_j n^\kappa \mathrm{E}\|\widehat{C} - C\|_{\mathcal{L}}^{\beta} \le A_j L_\beta(n) n^{-\beta(1-2/\alpha)+\kappa}.$$

Since $-\beta(1 - 2/\alpha) + \kappa < 0$, by Proposition 2.6 (i) of Resnick [31], we obtain (24).

Assumption 2 thus always holds as long as the eigenvalue separations d_j defined by (20) are positive, but this is a sufficient condition, so we state what is needed for our results to hold.

Assumption 2. The estimators $\hat{v}_j$ satisfy (22) if $\alpha > 4$ and (24) if $\alpha \in (2, 4)$.

Asymptotic properties in extreme value theory are typically derived as the number of upper order statistics, k, tends to infinity with the sample size n, in such a way that $k/n \to 0$. This condition remains to be sufficient for $\mathrm{D}_n(\xi_j, \xi_{j'}) \xrightarrow{P} \mathrm{D}(\xi_j, \xi_{j'})$, since the population scores $\mathbf{Y}_i$ are i.i.d. and regularly varying under Assumption 1. In our setting, however, we estimate the EDM based on $\widehat{\mathrm{D}}_n(\xi_j, \xi_{j'})$ calculated from the observed approximations $\widehat{\mathbf{Y}}_i$. It can be therefore expected that this additional approximation will, to some extent, restrict the rate at which k tends to infinity with n. We formulate a sufficient condition on the order of k in Assumption 3. We first define the function

$$\gamma(\alpha) = \begin{cases} \dfrac{6-\alpha}{\alpha+2}, & \alpha \in (2, 4], \\[2ex] \dfrac{\alpha-2}{2\alpha-2}, & \alpha \in (4, \infty). \end{cases} \tag{25}$$

Fig. 3 shows that $\gamma(\cdot)$ is continuous at the phase transition point $\alpha = 4$ with $\gamma(4) = 1/3$. It increases on $(4, \infty)$ with $\lim_{\alpha \nearrow \infty} \gamma(\alpha) = \frac{1}{2}$. For $\alpha \in (2, 4)$, $\gamma(\alpha)$ decreases with $\lim_{\alpha \searrow 2} \gamma(\alpha) = 1$. For each value of $\alpha > 2$, the interval $(\gamma(\alpha), 1)$ is not empty. We write

$$k >> n^\gamma, \quad \text{for some } \gamma \in (0, 1), \text{ if } k/n^\gamma \to \infty.$$

Assumption 3. We assume that $k \gg n^{\gamma}$ for some $\gamma \in (\gamma(\alpha), 1)$, with $\gamma(\alpha)$ defined in (25).

Assumption 3 implies that $k > \sqrt{n}$ always works if $\alpha > 4$, but as $\alpha \searrow 2$, almost all observations must be used to ensure the consistency of the estimator.

With all assumptions formulated and explained, we are ready to state the first main result of this section.

Theorem 1. *Recall the definitions of the EDM $\mathrm{D}(\xi_j, \xi_{j'})$ and its estimator $\widehat{\mathrm{D}}_n(\xi_j, \xi_{j'})$ given, respectively, in (13) and (14). Under Assumptions 1, 2, and 3,*

$$\widehat{\mathrm{D}}_n(\xi_j, \xi_{j'}) \xrightarrow{P} \mathrm{D}(\xi_j, \xi_{j'}).$$

Recall that $\mathrm{D}(\xi_j, \xi_{j'})$ integrates extremal dependence over the whole sphere $\mathbb{S}^2$, so we decompose $\mathrm{D}(\xi_j, \xi_{j'})$ into components measuring dependence over the four quadrants:

$$\begin{aligned}&\left[\mathrm{D}^{(+,+)}(\xi_j, \xi_{j'}), \mathrm{D}^{(-,+)}(\xi_j, \xi_{j'}), \mathrm{D}^{(-,-)}(\xi_j, \xi_{j'}), \mathrm{D}^{(+,-)}(\xi_j, \xi_{j'})\right]\\ &:= \left[\int_{\mathbb{S}^2_{(+,+)}} a_1 a_2 \Gamma_{jj'}(d\mathbf{a}),\ \int_{\mathbb{S}^2_{(-,+)}} a_1 a_2 \Gamma_{jj'}(d\mathbf{a}),\ \int_{\mathbb{S}^2_{(-,-)}} a_1 a_2 \Gamma_{jj'}(d\mathbf{a}),\ \int_{\mathbb{S}^2_{(+,-)}} a_1 a_2 \Gamma_{jj'}(d\mathbf{a})\right].\end{aligned}$$

The corresponding estimators for the components are given by, respectively,

$$\begin{aligned}\widehat{\mathrm{D}}_n^{(+,+)}(\xi_j, \xi_{j'}) &:= \frac{1}{k}\sum_{i=1}^n \frac{\hat{\xi}_{ij}}{\widehat{R}_i}\frac{\hat{\xi}_{ij'}}{\widehat{R}_i} I_{\widehat{R}_i \ge \widehat{R}_{(k)}} I_{\hat{\xi}_{ij}\ge 0, \hat{\xi}_{ij'}\ge 0}, \quad \widehat{\mathrm{D}}_n^{(-,+)}(\xi_j, \xi_{j'}) := \frac{1}{k}\sum_{i=1}^n \frac{\hat{\xi}_{ij}}{\widehat{R}_i}\frac{\hat{\xi}_{ij'}}{\widehat{R}_i} I_{\widehat{R}_i \ge \widehat{R}_{(k)}} I_{\hat{\xi}_{ij}\le 0, \hat{\xi}_{ij'}\ge 0},\\ \widehat{\mathrm{D}}_n^{(-,-)}(\xi_j, \xi_{j'}) &:= \frac{1}{k}\sum_{i=1}^n \frac{\hat{\xi}_{ij}}{\widehat{R}_i}\frac{\hat{\xi}_{ij'}}{\widehat{R}_i} I_{\widehat{R}_i \ge \widehat{R}_{(k)}} I_{\hat{\xi}_{ij}\le 0, \hat{\xi}_{ij'}\le 0}, \quad \widehat{\mathrm{D}}_n^{(+,-)}(\xi_j, \xi_{j'}) := \frac{1}{k}\sum_{i=1}^n \frac{\hat{\xi}_{ij}}{\widehat{R}_i}\frac{\hat{\xi}_{ij'}}{\widehat{R}_i} I_{\widehat{R}_i \ge \widehat{R}_{(k)}} I_{\hat{\xi}_{ij}\ge 0, \hat{\xi}_{ij'}\le 0}.\end{aligned} \tag{26}$$

Note that k in (26) is the same as in (14). In application, we first select k, the number of upper order statistics $\widehat{R}_{(i)}$ and then use it to compute (14) and (26). We will describe this with details in Section 2 of the supplement.

We establish the consistency of these estimators in the following corollary.

Corollary 2. *Under Assumptions 1, 2, and 3,*

$$\widehat{\mathrm{D}}_n^{(+,+)}(\xi_j, \xi_{j'}) \xrightarrow{P} \mathrm{D}^{(+,+)}(\xi_j, \xi_{j'}), \quad \widehat{\mathrm{D}}_n^{(-,+)}(\xi_j, \xi_{j'}) \xrightarrow{P} \mathrm{D}^{(-,+)}(\xi_j, \xi_{j'}),$$

$$\widehat{\mathrm{D}}_n^{(-,-)}(\xi_j, \xi_{j'}) \xrightarrow{P} \mathrm{D}^{(-,-)}(\xi_j, \xi_{j'}), \quad \widehat{\mathrm{D}}_n^{(+,-)}(\xi_j, \xi_{j'}) \xrightarrow{P} \mathrm{D}^{(+,-)}(\xi_j, \xi_{j'}).$$

Theorem 1 and Corollary 2 are proven in Section 5. Our approach to prove the consistency for the EDM is based on weak convergence of tail empirical measures. Set $\widehat{\boldsymbol{\Theta}}_i = \widehat{\mathbf{Y}}_i/\|\widehat{\mathbf{Y}}_i\|$. The estimator $\widehat{\mathrm{D}}_n(\xi_j, \xi_{j'})$ can then be written as an integral of a tail empirical measure, i.e.,

$$\widehat{\mathrm{D}}_n(\xi_j, \xi_{j'}) = \int_{\mathbb{S}^2} a_1 a_2 \widehat{\Gamma}_n(d\mathbf{a}), \quad \widehat{\Gamma}_n := \frac{1}{k}\sum_{i=1}^n I_{\widehat{\boldsymbol{\Theta}}_i} I_{\widehat{R}_i \ge \widehat{R}_{(k)}}.$$

The key argument to prove the consistency is therefore to show

$$\widehat{\Gamma}_n \Rightarrow \Gamma_{jj'} \quad \text{in } M_+\left(\mathbb{S}^2\right), \tag{27}$$

with $\Gamma_{jj'}$ in (13). Relation (27) is established by proving a series of weak convergence results.

We now turn to the asymptotic normality. The asymptotic normality of the estimator for the EDM is proven for i.i.d. bivariate observations in Larsson and Resnick [22]. To show the asymptotic normality of an estimator based on heavy-tailed data, additional conditions are required even in fully observable i.i.d. settings. For example, for the Hill estimator, second-order regular variation with restrictions on the rate of k is assumed, see Haeusler and Teugels [16], Csörgő et al. [4], Resnick and Stărică [32,33]. The aforementioned condition is a univariate concept, which is not applicable to our context. Instead, we use a multivariate version of second-order regular variation, defined by Resnick [29]. With some constraint on k, i.e., $\sqrt{k}A(b(n/k)) \to 0$, where A is defined in formula (15) in Resnick [29], the multivariate second-order regular variation implies the following weaker condition, which is also assumed by Larsson and Resnick [22].

Assumption 4. The R_i, $\boldsymbol{\Theta}_i$ satisfy

$$\sqrt{k}\left[\frac{n}{k}\Pr\left(\left(\frac{R_1}{b(n/k)}, \boldsymbol{\Theta}_1\right) \in \cdot\right) - c\nu_\alpha \times \Gamma_{jj'}\right] \xrightarrow{v} 0 \quad \text{in } M_+((0,\infty] \times \mathbb{S}^2).$$

Assumption 4 means that R_1 and $\boldsymbol{\Theta}_1$ are asymptotically independent. We emphasize that this assumption applies to population quantities, which are not observable in our setting. We now formulate the asymptotic normality of our estimator for the EDM, which is based on projections of functional data.

Theorem 2. *Under Assumptions 2, 3 and 4*

$$\sqrt{k}\left(\widehat{\mathrm{D}}_n(\xi_j, \xi_{j'}) - \mathrm{D}(\xi_j, \xi_{j'})\right) \Rightarrow \mathcal{N}(0, \sigma^2),$$

where $\sigma^2 = \mathrm{Var}(\widetilde{\Theta}_1\widetilde{\Theta}_2) > 0$, with $\widetilde{\Theta}_1$ and $\widetilde{\Theta}_2$ being the components of a random vector with distribution $\Gamma_{jj'}$.

4. Preliminary results

We put together several preliminary results in this section to avoid burdening the proofs in Section 5, so that readers can keep track of the main flow of the argument made in Section 5.

The first lemma follows from Lemma 3.7 of Kim and Kokoszka [20] and is needed to prove Lemma 2.

Lemma 1. *Suppose random variables $H_m(n)$, $m, n \geq 1$, satisfy $0 \leq H_m(n) \leq 1$ and $\forall\, m \geq 1$, $\quad H_m(n) \xrightarrow{P} 0, \quad$ as $n \to \infty$. Then, $\sum_{m=1}^{\infty} 2^{-m} H_m(n) \xrightarrow{P} 0, \quad$ as $n \to \infty$.*

In the following lemma, we present a sufficient condition to guarantee the convergence between random measures defined on a nice space. We denote a locally compact topological space with countable base by $\mathbb{E}$. Following page 51 of Resnick [31], the vague metric $d(\cdot, \cdot)$ on $M_+(\mathbb{E})$ is defined by

$$d(\mu_1, \mu_2) = \sum_{i=1}^{\infty} \frac{|\mu_1(f_i) - \mu_2(f_i)| \wedge 1}{2^i}, \qquad \mu_1, \mu_2 \in M_+(\mathbb{E}), \tag{28}$$

for some sequence of functions $f_i \in C_K^+(\mathbb{E})$ where $C_K^+(\mathbb{E})$ is the space of continuous functions with compact support on $\mathbb{E}$. By Lemma 1, the following is readily proven.

Lemma 2. *Suppose that μ_n, ν_n are random measures in $M_+(\mathbb{E})$. If, for any $f \in C_K^+(\mathbb{E})$, $|\mu_n(f) - \nu_n(f)| \xrightarrow{P} 0, \quad n \to \infty$, then $d(\mu_n, \nu_n) \xrightarrow{P} 0$.*

In the following lemma, we show that a continuous mapping with a compactness condition preserves convergence of random measures. Suppose that $\mathbb{E}_1$ and $\mathbb{E}_2$ are locally compact topological spaces with countable base. Denote by $\mathcal{K}(\mathbb{E})$ a set of all compact subsets of $\mathbb{E}$.

Lemma 3. *Suppose that $H : \mathbb{E}_1 \mapsto \mathbb{E}_2$ is a continuous function such that*

$$H^{-1}(K_2) \in \mathcal{K}(\mathbb{E}_1), \quad \forall K_2 \in \mathcal{K}(\mathbb{E}_2). \tag{29}$$

If random measures μ_n, ν_n in $M_+(\mathbb{E}_1)$ satisfy $d(\mu_n, \nu_n) \xrightarrow{P} 0$, as $n \to \infty$, then $d(\mu_n \circ H^{-1}, \nu_n \circ H^{-1}) \xrightarrow{P} 0$, in $M_+(\mathbb{E}_2)$.

Proof. By Lemma 2, it suffices to show that, for any $f \in C_K^+(\mathbb{E}_2)$,

$$\mu_n \circ H^{-1}(f) - \nu_n \circ H^{-1}(f) \xrightarrow{P} 0. \tag{30}$$

Using the change of variables, we have $(\mu_n - \nu_n) \circ H^{-1}(f) = \int_{\mathbb{E}_2} f(e_2)(\mu_n - \nu_n) \circ H^{-1}(de_2) = \int_{\mathbb{E}_1} f(H(e_1))(\mu_n - \nu_n)(de_1)$. Thus, we have $(\mu_n - \nu_n) \circ H^{-1}(f) = (\mu_n - \nu_n)(f \circ H)$. Since f and H are both continuous, and with (29), we get $f \circ H \in C_K^+(\mathbb{E}_1)$, see page 142 of Resnick [31]. Then, since $d(\mu_n, \nu_n) \xrightarrow{P} 0$ by assumption, we get (30). □

Consider the polar coordinate transform $T : [-\infty, \infty]^2 \setminus \{\mathbf{0}\} \mapsto (0, \infty] \times \mathbb{S}^2$ defined by, for $\mathbf{x} \in [-\infty, \infty]^2 \setminus \{\mathbf{0}\}$,

$$T(\mathbf{x}) = \left(\|\mathbf{x}\|, \frac{\mathbf{x}}{\|\mathbf{x}\|}\right). \tag{31}$$

Note that T is not bijective since its boundaries at infinity are included. Thus, Lemma 3 cannot be directly applied to T to show that it preserves convergence of random measures. Instead, we will show that by using, say, "restrict and then extend space" strategy, which is used in a different setting on page 176∼179 of Resnick [31]. We follow the technique in the proof of the next lemma.

Lemma 4. *Suppose that random measures μ_n, ν_n satisfy*

$$d(\mu_n, \nu_n) \xrightarrow{P} 0, \quad \text{in } M_+([-\infty, \infty]^2 \setminus \{\mathbf{0}\}), \tag{32}$$

as $n \to \infty$. Then, $d(\mu_n \circ T^{-1}, \nu_n \circ T^{-1}) \xrightarrow{P} 0, \quad$ in $M_+((0, \infty] \times \mathbb{S}^2)$.

Proof. Consider the transform $T' : (-\infty, \infty)^2 \setminus \{\mathbf{0}\} \mapsto (0, \infty) \times \mathbb{S}^2$ defined by (31). Our first claim is that (32) implies

$$d(\mu_n, \nu_n) \xrightarrow{P} 0, \quad \text{in } M_+((-\infty, \infty)^2 \setminus \{\mathbf{0}\}). \tag{33}$$

Let $f_i \in C_K^+ \left((-\infty, \infty)^2 \setminus \{\mathbf{0}\}\right)$, and suppose that $K_i \in \mathcal{K}((-\infty, \infty)^2 \setminus \{\mathbf{0}\})$ is the compact support of f_i. Let $\tilde{f}_i := f_i(x) I_{x \in K_i}$, then $\tilde{f}_i \in C_K^+ \left([-\infty, \infty]^2 \setminus \{\mathbf{0}\}\right)$. Observe that $d(\mu_n, \nu_n) = \sum_{i=1}^{\infty} 2^{-i} |(\mu_n - \nu_n)(\tilde{f}_i)| = \sum_{i=1}^{\infty} 2^{-i} |(\mu_n - \nu_n)(f_i)| \xrightarrow{P} 0$, by (32), so we get (33).

Our second claim is that (33) implies

$$d(\mu_n \circ (T')^{-1}, \nu_n \circ (T')^{-1}) \xrightarrow{P} 0, \quad \text{in } M_+((0, \infty) \times \mathbb{S}^2). \tag{34}$$

This is readily proven by Lemma 3, since T' is continuous and satisfy (29).

The last step is now to extend T' to the bigger space, where ∞ is included. Let $f_i \in C_K^+ \left((0, \infty] \times \mathbb{S}^2\right)$, and set $\|f_i\| = \sup f_i < \infty$. We define a smooth truncation function of r, for fixed M, δ, by

$$\phi(r; M, \delta) := I_{0 < r \le M} + \{-(r - M)/\delta + 1\} I_{M < r \le M + \delta}.$$

Then, observe that

$$d(\mu_n \circ T^{-1}, \nu_n \circ T^{-1}) = \sum_{i=1}^{\infty} 2^{-i} |(\mu_n - \nu_n) \circ T^{-1}(f_i)| - \sum_{i=1}^{\infty} 2^{-i} |(\mu_n - \nu_n) \circ T^{-1}(f_i \phi)|$$
$$+ \sum_{i=1}^{\infty} 2^{-i} |(\mu_n - \nu_n) \circ T^{-1}(f_i \phi)| - \sum_{i=1}^{\infty} 2^{-i} |(\mu_n - \nu_n) \circ (T')^{-1}(f_i \phi)| + \sum_{i=1}^{\infty} 2^{-i} |(\mu_n - \nu_n) \circ (T')^{-1}(f_i \phi)| =: A + B + C.$$

Now, we will show that each of the components goes to 0. First, observe that

$$A \le \sum_{i=1}^{\infty} 2^{-i} \left| \int_{(0,\infty] \times \mathbb{S}^2} f_i(r, \theta)(1 - \phi(r))(\mu_n - \nu_n) \circ T^{-1}(dr, d\theta) \right| \le \sum_{i=1}^{\infty} 2^{-i} \|f_i\| \left| \int_{(M,\infty] \times \mathbb{S}^2} (\mu_n - \nu_n) \circ T^{-1}(dr, d\theta) \right|.$$

Taking a sufficiently large M, then A gets arbitrarily small. Next, for each M,

$$B \le \sum_{i=1}^{\infty} 2^{-i} \|f_i\| \left| \int_{(0,M] \times \mathbb{S}^2} (\mu_n - \nu_n) \circ (T^{-1} - (T')^{-1})(dr, d\theta) \right| = 0.$$

Since $f_i(r, \theta)\phi(r; M, \delta) \in C_K^+ \left((0, \infty) \times \mathbb{S}^2\right)$, the last term C goes to 0 by (34). □

The next lemma shows that the distance between a population score and its corresponding approximation is asymptotically negligible.

Lemma 5. *Under Assumptions 1, 2, for $\alpha > 4$, $\mathrm{E}|\hat{\xi}_j - \xi_j| = O(n^{-1/2})$, and for $2 < \alpha < 4$, $\mathrm{E}|\hat{\xi}_j - \xi_j|^r = o(n^{-\kappa r/\beta})$, for some $r > 0$ satisfying*

$$r < \frac{2\beta}{\beta + 2}, \tag{35}$$

where κ, β are defined in (24).

Proof. For $\alpha > 4$, by the Cauchy–Schwarz inequality, $|\hat{\xi}_j - \xi_j| \le \|X\| \|\hat{v}_j - v_j\|$, so by Assumption 2,

$$\mathrm{E}|\hat{\xi}_j - \xi_j| \le \{\mathrm{E}\|X\|^2\}^{1/2} \{\mathrm{E}\|\hat{v}_j - v_j\|^2\}^{1/2} = O(n^{-1/2}).$$

Now consider the case of $2 < \alpha < 4$. Since for any β, $\frac{2\beta}{\beta+2} < \beta$, condition (35), implies that $r < \beta$. Applying Hölder's inequality with $p = \beta/r > 1$ and $q = \beta/(\beta - r)$, we get $\mathrm{E}|\hat{\xi}_j - \xi_j|^r \le \{\mathrm{E}\|\hat{v}_j - v_j\|^\beta\}^{\frac{r}{\beta}} \{\mathrm{E}\|X\|^{rq}\}^{1/q}$. Direct verification shows that condition (35) is equivalent to

$$\frac{2\beta^2}{(\beta + 2)(\beta - r)} < 2,$$

which implies $rq < 2$. Hence, by Assumption 1, $\{\mathrm{E}\|X\|^{rq}\}^{1/q} < \infty$. Therefore, by (24), $\mathrm{E}|\hat{\xi}_j - \xi_j|^r = o(n^{-\kappa r/\beta})$. □

In the following lemmas, we verify the continuity of functions that will be used in Section 5 with the continuous mapping theorem.

Lemma 6. *Suppose that the map $H : M_+ \left((0, \infty] \times \mathbb{S}^2\right) \times (0, \infty) \to M_+((0, \infty] \times \mathbb{S}^2)$, defined by for any measurable set $A \times B \subset (0, \infty] \times \mathbb{S}^2$,*

$$H(U, x)(A \times B) = U(xA \times B).$$

The map H is continuous at $(\nu_\alpha \times \Gamma_{jj'}, x)$.

Proof. Suppose $W_n \xrightarrow{v} \nu_\alpha \times \Gamma_{jj'}$ in $M_+\left((0,\infty]\times\mathbb{S}^2\right)$, and $x_n \to x$ in $(0,\infty)$. Then we must show that

$$H(W_n, x_n) = W_n((x_n\cdot)\times\cdot) \xrightarrow{v} H(\nu_\alpha \times \Gamma_{jj'}, x) = \nu_\alpha \times \Gamma_{jj'}((x\cdot)\times\cdot).$$

To verify this, it suffices to show that for any $f \in C_K^+((0,\infty]\times\mathbb{S}^2)$,

$$W_n((x_n\cdot)\times\cdot)(f) = \int_{(0,\infty]\times\mathbb{S}^2} f(t,\mathbf{a})\, W_n(x_n dt, d\mathbf{a}) = \int_{(0,\infty]\times\mathbb{S}^2} f(y/x_n,\mathbf{a})\, W_n(dy, d\mathbf{a})$$
$$\to \nu_\alpha \times \Gamma_{jj'}((x\cdot)\times\cdot)(f) = \int_{(0,\infty]\times\mathbb{S}^2} f(t,\mathbf{a})\, \nu_\alpha(x_n dt)\Gamma_{jj'}(d\mathbf{a}) = \int_{(0,\infty]\times\mathbb{S}^2} f(y/x,\mathbf{a})\, \nu_\alpha(dy)\Gamma_{jj'}(d\mathbf{a}).$$

The following verification is mostly based on pp. 83–84 of Resnick [31], whose test functions are univariate. Our test functions are however bivariate. We must employ a product metric to apply uniform continuity of the test functions.

First observe that

$$\left|\int_{(0,\infty]\times\mathbb{S}^2_+} f(y/x_n,\mathbf{a})\, W_n(dy,d\mathbf{a}) - \int_{(0,\infty]\times\mathbb{S}^2} f(y/x,\mathbf{a})\, \nu_\alpha(dy)\Gamma_{jj'}(d\mathbf{a})\right|$$
$$\le \left|\int_{(0,\infty]\times\mathbb{S}^2} f(y/x_n,\mathbf{a})\, W_n(dy,d\mathbf{a}) - \int_{(0,\infty]\times\mathbb{S}^2} f(y/x,\mathbf{a})\, W_n(dy,d\mathbf{a})\right|$$
$$+ \left|\int_{(0,\infty]\times\mathbb{S}^2} f(y/x,\mathbf{a})\, W_n(dy,d\mathbf{a}) - \int_{(0,\infty]\times\mathbb{S}^2} f(y/x,\mathbf{a})\, \nu_\alpha(dy)\Gamma_{jj'}(d\mathbf{a})\right|.$$

Since $W_n \xrightarrow{v} \nu_\alpha \times \Gamma_{jj'}$ and $f(\frac{\cdot}{x},\cdot) \in C_K^+((0,\infty]\times\mathbb{S}^2)$, the second term of the right-hand side goes to zero. Now, we focus on the first term. Since f has compact support in $(0,\infty]\times\mathbb{S}^2$, we can take $\delta > 0$ such that the supports of $f(\frac{\cdot}{x},\cdot)$ and $f(\frac{\cdot}{x_n},\cdot)$, for large n, are contained in $[\delta,\infty]\times\mathbb{S}^2$. Then we get the bound

$$\left|\int_{(0,\infty]\times\mathbb{S}^2_+} f(y/x_n,\mathbf{a})\, W_n(dy,d\mathbf{a}) - \int_{(0,\infty]\times\mathbb{S}^2_+} f(y/x,\mathbf{a})\, W_n(dy,d\mathbf{a})\right|$$
$$\le \int_{[\delta,\infty]\times\mathbb{S}^2_+} |f(y/x_n,\mathbf{a}) - f(y/x,\mathbf{a})|\, W_n(dy,d\mathbf{a}) \le \sup_{y\ge\delta,\ \mathbf{a}\in\mathbb{S}^2} |f(y/x_n,\mathbf{a}) - f(y/x,\mathbf{a})|\, W_n([\delta,\infty]\times\mathbb{S}^2).$$

Since $W_n([\delta,\infty]\times\mathbb{S}^2)$ is bounded, it remains to show that as $x_n \to x$,

$$\sup_{y\ge\delta,\ \mathbf{a}\in\mathbb{S}^2} |f(y/x_n,\mathbf{a}) - f(y/x,\mathbf{a})| \to 0. \tag{36}$$

We use the fact that a continuous function with compact support is uniformly continuous. The metric on $(0,\infty]\times\mathbb{S}^2$ is given by $d_{\text{prod}}((u,\mathbf{a}),(v,\mathbf{b})) = d_{(0,\infty]}(u,v) + d_{\mathbb{S}^2}(\mathbf{a},\mathbf{b})$, see p.57 of Resnick [31]. Define the metric on $(0,\infty]$ by

$$d_{(0,\infty]}(u,v) = |u^{-1} - v^{-1}|,$$

for $u, v \in (0,\infty]$, which measures the distance between points in $(0,\infty]$ with one point compactification at ∞. Since $x_n \to x$ and $y \ge \delta_0$,

$$d_{\text{prod}}\left((y/x_n,\mathbf{a}),(y/x,\mathbf{a})\right) = \frac{|x_n - x|}{y} \le \frac{|x_n - x|}{\delta_0} \to 0.$$

Therefore, by the uniform continuity of f, we get (36). □

Lemma 7. *The function g on $M_+\left((0,\infty]\times\mathbb{S}^2\right)$ defined by for any measurable sets $A \subset (0,\infty]$, $B \subset \mathbb{S}^2$, $g(U) = U(A\times B)$ is continuous at $\nu_\alpha \times \Gamma_{jj'}$.*

Proof. Suppose $W_n \xrightarrow{v} \nu_\alpha \times \Gamma_{jj'}$ in $M_+\left((0,\infty]\times\mathbb{S}^2\right)$. Since $A\times B$ is relatively compact in $(0,\infty]\times\mathbb{S}^2$, by Theorem 3.2 of Resnick [31] $g(W_n) = W_n(A\times B) \to g(\nu_\alpha\times\Gamma_{jj'}) = \nu_\alpha(A)\Gamma_{jj'}(B)$. □

Lemma 8. *The function h on $M_+\left(\mathbb{S}^2\right)$ defined by for $B \in \{\mathbb{S}^2, \mathbb{S}^2_{(+,+)}, \mathbb{S}^2_{(-,+)}, \mathbb{S}^2_{(-,-)}, \mathbb{S}^2_{(+,-)}\}$, $h(U) = \int_B \theta_1\theta_2 U(d\boldsymbol{\theta})$ is continuous at $\Gamma_{jj'}$.*

Proof. Suppose $W_n \xrightarrow{v} \Gamma_{jj'}$ in $M_+\left(\mathbb{S}^2\right)$. Consider a map $f : \mathbb{S}^2 \to \mathbb{R}$, defined by $f(\boldsymbol{\theta}) = \theta_1\theta_2 I_{\boldsymbol{\theta}\in B}$. Note that every continuous function on a compact space has compact support. Since f is continuous with compact support, by the definition of vague convergence,

$$h(W_n) = \int_B \theta_1\theta_2 W_n(d\boldsymbol{\theta}) \to h(\Gamma_{jj'}) = \int_B \theta_1\theta_2 \Gamma_{jj'}(d\boldsymbol{\theta}). \quad \square$$

5. Proofs of the results of Section 3

Proof of Proposition 3. First, note that $\|\pi(X)\| > yb(n)$ and $\pi(X)/\|\pi(X)\| \in \cdot$ iff $(yb(n))^{-1}X \in \mathcal{A}_\pi(\cdot)$. Observe that, for any set S in $\mathcal{B}(\mathbb{S}^d)$,

$$n\mathrm{Pr}(\|\pi(X)\| > yb(n),\ \pi(X)/\|\pi(X)\| \in S) = n\,\mathrm{Pr}\left(\frac{X}{yb(n)} \in \mathcal{A}_\pi(S)\right).$$

To prove the regular variation of $\pi(X)$ in $\mathbb{R}^d$, we will apply Theorem 2.3 of Lindskog et al. [26]. To do this, we must show that the $\mathcal{A}_\pi(S)$ are continuity sets of ν, i.e., $\nu(\partial\mathcal{A}_\pi(S)) = 0$. The verification uses the same idea described in the proof of Proposition 3.1 of Kokoszka et al. [21], but the difference is that we work with the different projection $\pi(z)$ and its relevant set $\mathcal{A}_\pi(S)$.

By (18), we have

$$\partial\mathcal{A}_\pi(S) = \{z \in \mathbb{H} : \|\pi(z)\| = 1,\ \pi(z)/\|\pi(z)\| \in S\},\ \partial(r\mathcal{A}_\pi(S)) = \{z \in \mathbb{H} : \|\pi(z)\| = r,\ \pi(z)/\|\pi(z)\| \in S\}.$$

Note that $\partial(r\mathcal{A}_\pi(S)) = r\partial\mathcal{A}_\pi(S)$, and the sets $\partial(r\mathcal{A}_\pi(S))$ are all disjoint in r. We assume $\nu(\partial\mathcal{A}_\pi(S)) > 0$ and get a contradiction. Since $\mathcal{A}_\pi(S) \supset \cup_{n\ge 1}\partial(n^{1/\alpha}\mathcal{A}_\pi(S))$, for all $\alpha > 0$, and ν is homogeneous,

$$\nu(\mathcal{A}_\pi(S)) \ge \sum_{n=1}^{\infty} \nu(n^{1/\alpha}\partial\mathcal{A}_\pi(S)) = \sum_{n=1}^{\infty} n^{-1}\nu(\partial\mathcal{A}_\pi(S)) = \infty.$$

This contradicts to the fact that ν is boundedly finite. Therefore, the $\mathcal{A}_\pi(S)$ are continuity sets of ν.

Now, by Theorem 2.3 of Lindskog et al. [26] and (15), we obtain

$$n\mathrm{Pr}(\|\pi(X)\| > yb(n),\ \pi(X)/\|\pi(X)\| \in S) \to \nu(y\mathcal{A}_\pi(S)) = y^{-\alpha}\nu(\mathcal{A}_\pi(S)).$$

Setting

$$\Gamma(\cdot) := \frac{\nu(\mathcal{A}_\pi(\cdot))}{c}, \qquad c = \nu(\mathcal{A}_\pi(\mathbb{S}^d)), \tag{37}$$

we get the claim.

Proof of Theorem 1. Recall that

$$\mathbf{Y}_i = [\xi_{ij}, \xi_{ij'}]^\top,\ R_i = \|\mathbf{Y}_i\|,\ \boldsymbol{\Theta}_i = \mathbf{Y}_i/R_i,\ \widehat{\mathbf{Y}}_i = [\hat{\xi}_{ij}, \hat{\xi}_{ij'}]^\top,\ \widehat{R}_i = \|\widehat{\mathbf{Y}}_i\|,\ \widehat{\boldsymbol{\Theta}}_i = \widehat{\mathbf{Y}}_i/\widehat{R}_i.$$

Under Assumption 1, the $\mathbf{Y}_i$ are regularly varying with index $-\alpha$ by Proposition 3. More specifically, there exist a sequence $\{b(n)\}$ (the same as in (15)) and a probability angular measure $\Gamma_{jj'}$ defined as (37) satisfying

$$n\mathrm{Pr}\left(\left(\frac{R_i}{b(n)},\ \boldsymbol{\Theta}_i\right) \in \cdot\right) \xrightarrow{v} c\nu_\alpha \times \Gamma_{jj'} \quad \text{in } M_+((0,\infty] \times \mathbb{S}^2). \tag{38}$$

The constant c depends on the choice of $b(n)$. In the following, we assume $c = 1$ to keep the notation simple.

Our approach is to establish several weak convergences of tail empirical measures. We start with an empirical measure based on i.i.d. $\mathbf{Y}_i$:

$$U_n := \frac{1}{k}\sum_{i=1}^{n} I_{(R_i/b(n/k),\ \boldsymbol{\Theta}_i)} \Rightarrow \nu_\alpha \times \Gamma_{jj'} \quad \text{in } M_+\left((0,\infty] \times \mathbb{S}^2\right). \tag{39}$$

We then extend (39) to

$$\widehat{U}_n := \frac{1}{k}\sum_{i=1}^{n} I_{\left(\widehat{R}_i/b(n/k),\ \widehat{\boldsymbol{\Theta}}_i\right)} \Rightarrow \nu_\alpha \times \Gamma_{jj'} \quad \text{in } M_+\left((0,\infty] \times \mathbb{S}^2\right). \tag{40}$$

Since the $\widehat{\mathbf{Y}}_i$ are no longer independent, this requires techniques involving the Slutsky theorem. We further proceed to replace the unknown sequence $b(n/k)$ by its estimate $\widehat{R}_{(k)}$:

$$\widehat{U}_n^\star := \frac{1}{k}\sum_{i=1}^{n} I_{\left(\widehat{R}_i/\widehat{R}_{(k)},\ \widehat{\boldsymbol{\Theta}}_i\right)} \Rightarrow \nu_\alpha \times \Gamma_{jj'} \quad \text{in } M_+\left((0,\infty] \times \mathbb{S}^2\right). \tag{41}$$

Applying the continuous mapping theorem, we finally get (27), i.e.,

$$\widehat{\Gamma}_n = \frac{1}{k}\sum_{i=1}^{n} I_{\widehat{\boldsymbol{\Theta}}_i} I_{\widehat{R}_i \ge \widehat{R}_{(k)}} \Rightarrow \Gamma_{jj'} \quad \text{in } M_+\left(\mathbb{S}^2\right).$$

The consistency of $\widehat{\mathrm{D}}_n(\xi_j, \xi_{j'})$ for $\mathrm{D}(\xi_j, \xi_{j'})$ is then established because $\widehat{\mathrm{D}}_n(\xi_j, \xi_{j'}) = \int_{\mathbb{S}^2} a_1 a_2 \widehat{\Gamma}_n(d\mathbf{a})$.

We now present a series of the results mentioned above, of which Proposition 4 is the most essential and important step toward Theorem 1. The following lemma verifies (39), which is readily proven from (38) by Theorem 5.3 (ii) of Resnick [31].

Lemma 9. *Under Assumption 1, relation (39) holds.*

The next result shows that the infeasible samples $\mathbf{Y}_i$ in (39) can be replaced by their approximations $\widehat{\mathbf{Y}}_i$.

Proposition 4. *Under Assumptions 1, 2, and 3, relation (40) holds.*

Proof. By Lemma 9 and the Slutsky theorem, it suffices to prove that

$$d(\widehat{U}_n, U_n) = d\left(\frac{1}{k}\sum_{i=1}^{n} I_{\left(\widehat{R}_i/b(n/k),\ \widehat{\Theta}_i\right)}, \frac{1}{k}\sum_{i=1}^{n} I_{(R_i/b(n/k),\ \Theta_i)}\right) \xrightarrow{P} 0. \tag{42}$$

To show (42), we set $\widehat{V}_n := \frac{1}{k}\sum_{i=1}^{n} I_{\widehat{\mathbf{Y}}_i/b(n/k)}, \quad V_n := \frac{1}{k}\sum_{i=1}^{n} I_{\mathbf{Y}_i/b(n/k)}$, and prove

$$d(\widehat{V}_n, V_n) = d\left(\frac{1}{k}\sum_{i=1}^{n} I_{\widehat{\mathbf{Y}}_i/b(n/k)}, \frac{1}{k}\sum_{i=1}^{n} I_{\mathbf{Y}_i/b(n/k)}\right) \xrightarrow{P} 0. \tag{43}$$

Applying the polar transformation defined in (31), we get (42) from (43) by Lemma 4.

To prove (43), it suffices to show that, by Lemma 2, for any $f \in C_K^+([-\infty,\infty]^2 \setminus \{\mathbf{0}\})$, and any $\tau > 0$,

$$\Pr\left(\left|\frac{1}{k}\sum_{i=1}^{n} f\left(\frac{\widehat{\mathbf{Y}}_i}{b(n/k)}\right) - \frac{1}{k}\sum_{i=1}^{n} f\left(\frac{\mathbf{Y}_i}{b(n/k)}\right)\right| > \tau\right) \to 0. \tag{44}$$

Since f has compact support in $[-\infty,\infty]^2 \setminus \{\mathbf{0}\}$, set

$$a := \inf\{\|s\| : s \in \operatorname{supp}(f)\} > 0. \tag{45}$$

To prove (44), we consider a decomposition using the following sets. For $0 < \eta < a/2$, set

$$A_n(k) := \left\{1 \le i \le n : \left\|\frac{\widehat{\mathbf{Y}}_i}{b(n/k)} - \frac{\mathbf{Y}_i}{b(n/k)}\right\| \le \eta,\ \left\|\frac{\mathbf{Y}_i}{b(n/k)}\right\| \ge a - \eta\right\},$$

$$B_n(k) := \left\{1 \le i \le n : \left\|\frac{\widehat{\mathbf{Y}}_i}{b(n/k)} - \frac{\mathbf{Y}_i}{b(n/k)}\right\| \le \eta,\ \left\|\frac{\mathbf{Y}_i}{b(n/k)}\right\| < a - \eta\right\},$$

$$C_n(k) := \left\{1 \le i \le n : \left\|\frac{\widehat{\mathbf{Y}}_i}{b(n/k)} - \frac{\mathbf{Y}_i}{b(n/k)}\right\| > \eta\right\}.$$

Then, we have

$$\Pr\left(\left|\frac{1}{k}\sum_{i=1}^{n} f\left(\frac{\widehat{\mathbf{Y}}_i}{b(n/k)}\right) - \frac{1}{k}\sum_{i=1}^{n} f\left(\frac{\mathbf{Y}_i}{b(n/k)}\right)\right| > \tau\right) \le \Pr(S(A_n) > \tau/3) + \Pr(S(B_n) > \tau/3) + \Pr(S(C_n) > \tau/3),$$

where

$$S(A_n) = \frac{1}{k}\sum_{i \in A_n(k)} \left|f\left(\frac{\widehat{\mathbf{Y}}_i}{b(n/k)}\right) - f\left(\frac{\mathbf{Y}_i}{b(n/k)}\right)\right|,$$

and $S(B_n)$ and $S(C_n)$ are defined analogously with $\sum_{i \in B_n(k)}$ and $\sum_{i \in C_n(k)}$, respectively.

We will show that each of the three parts goes to 0. We first investigate $\Pr(S(A_n) > \tau/3)$. Since f is uniformly continuous,

$$w_\eta(f) := \sup_{\|\mathbf{x}-\mathbf{y}\| \le \eta,\ \mathbf{x},\mathbf{y} \in [-\infty,\infty]^2 \setminus \{\mathbf{0}\}} |f(\mathbf{x}) - f(\mathbf{y})| \to 0, \ \ \eta \to 0.$$

Note that

$$S(A_n) \le w_\eta(f)\frac{1}{k}\#\left\{1 \le i \le n : \left\|\frac{\mathbf{Y}_i}{b(n/k)}\right\| \ge a - \eta\right\} = w_\eta(f) U_n(E_{a-\eta}),$$

with the measure U_n defined in (39), and with the set $E_b \subset (0,\infty] \times \mathbb{S}^2$ defined by

$$E_b = \left\{(r,\theta) \in (0,\infty] \times \mathbb{S}^2 : r \ge b\right\}, \quad b > 0.$$

Now consider the function g on $M_+\left((0,\infty]\times\mathbb{S}^2\right)$, defined by, for any measurable set $A\subset(0,\infty]$, $g(U)=U\left(A\times\mathbb{S}^2\right)$. Then, by Lemma 7 and the continuous mapping theorem, for a fixed η, $U_n(E_{a-\eta})\xrightarrow{P}\nu_\alpha(a-\eta,\infty]=(a-\eta)^{-\alpha}$. Therefore,

$$\limsup_{n\to\infty}\Pr(S(A_n)>\tau/3)\le\Pr\left(w_\eta(f)(a-\eta)^{-\alpha}>\tau/3\right)\le\Pr\left(w_\eta(f)>2^{-\alpha}a^\alpha\tau/3\right).$$

By taking sufficiently small η, we can ensure that $\Pr\left(w_\eta(f)>2^{-\alpha}a^\alpha\tau/3\right)=0$, hence $\lim_{n\to\infty}\Pr(S(A_n)>\tau/3)=0$.

Next, we consider the second probability in the decomposition. Observe that for each $i\in B_n(k)$,

$$\left\|\frac{\widehat{\mathbf{Y}}_i}{b(n/k)}\right\|\le\left\|\frac{\widehat{\mathbf{Y}}_i}{b(n/k)}-\frac{\mathbf{Y}_i}{b(n/k)}\right\|+\left\|\frac{\mathbf{Y}_i}{b(n/k)}\right\|<a,\quad\left\|\frac{\mathbf{Y}_i}{b(n/k)}\right\|<a-\eta.$$

Thus, the two points $\widehat{\mathbf{Y}}_i/b(n/k),\mathbf{Y}_i/b(n/k)$ are outside of the support of f for all $i\in B_n(k)$, so $S(B_n)=0$ by construction, and so $\Pr(S(B_n)>\tau/3)=0$.

It remains to show that for any $\eta>0$, $\lim_{n\to\infty}\Pr(S(C_n)>\tau/3)=0$. Set

$$\|f\|_\infty=\sup_{\mathbf{x}\in[-\infty,\infty]^2\setminus\{\mathbf{0}\}}|f(\mathbf{x})|. \tag{46}$$

First, consider the case of $\alpha>4$. By Markov's inequality,

$$\begin{aligned}\Pr(S(C_n)>\tau/3)&\le\frac{3}{\tau k}\mathrm{E}\left[\sum_{i\in C_n(k)}\left|f\left(\frac{\widehat{\mathbf{Y}}_i}{b(n/k)}\right)-f\left(\frac{\mathbf{Y}_i}{b(n/k)}\right)\right|\right]\le\frac{6\|f\|_\infty}{\tau k}\mathrm{E}\left[\sum_{i=1}^n I_{\|\widehat{\mathbf{Y}}_i-\mathbf{Y}_i\|>\eta b(n/k)}\right]\\&\le\frac{6\|f\|_\infty}{\tau}\frac{n}{k}\Pr\left(\|\widehat{\mathbf{Y}}_i-\mathbf{Y}_i\|>\eta b(n/k)\right)\le\frac{6\|f\|_\infty}{\tau\eta}\frac{n}{kb(n/k)}\mathrm{E}\|\widehat{\mathbf{Y}}_i-\mathbf{Y}_i\|.\end{aligned}$$

Since all norms in $\mathbb{R}^2$ are equivalent, we get

$$\|\widehat{\mathbf{Y}}_i-\mathbf{Y}_i\|\le C\left(|\hat{\xi}_{ij}-\xi_{ij}|+|\hat{\xi}_{ij'}-\xi_{ij'}|\right), \tag{47}$$

for some $C>0$. Since $\mathrm{E}\|\widehat{\mathbf{Y}}_i-\mathbf{Y}_i\|\le O(n^{-1/2})$ by Lemma 5, we have $\Pr(S(C_n)>\tau/3)=O\left(n^{1/2}/\{kb(n/k)\}\right)$. By Assumption 3 and (17), $\Pr(S(C_n)>\tau/3)=o(1)$.

Now consider the case of $\alpha\in(2,4)$. We will use Lemma 5, which refers to relation (24). Observe that since $\beta<\alpha/2<2$ in (24), it holds that $\frac{2\beta}{\beta+2}<1$. This implies that r satisfying (35) also satisfies $r<1$. Applying Markov's and Lyapunov's inequalities, we thus obtain

$$\begin{aligned}\Pr(S(C_n)>\tau/3)&\le\Pr\left(\frac{2\|f\|_\infty}{k}\sum_{i=1}^n I_{\|\widehat{\mathbf{Y}}_i-\mathbf{Y}_i\|>\eta b(n/k)}>\frac{\tau}{3}\right)\le\frac{6^r\|f\|_\infty^r n^r}{\tau^r k^r}\mathrm{E}\left[\left(\frac{1}{n}\sum_{i=1}^n I_{\|\widehat{\mathbf{Y}}_i-\mathbf{Y}_i\|>\eta b(n/k)}\right)^r\right]\\&\le\frac{6^r\|f\|_\infty^r n^r}{\tau^r k^r}\left\{\mathrm{E}\left[\frac{1}{n}\sum_{i=1}^n I_{\|\widehat{\mathbf{Y}}_i-\mathbf{Y}_i\|>\eta b(n/k)}\right]\right\}^r=\frac{6^r\|f\|_\infty^r n^r}{\tau^r k^r}\Pr\left(\|\widehat{\mathbf{Y}}_i-\mathbf{Y}_i\|>\eta b(n/k)\right)^r.\end{aligned}$$

Applying Markov's inequality with the same r again and (47), we obtain

$$\Pr(S(C_n)>\tau/3)\le c\frac{n^r}{k^r\{b(n/k)\}^{r^2}}\left\{\mathrm{E}\left[\max\left(|\hat{\xi}_{ij}-\xi_{ij}|,|\hat{\xi}_{ij'}-\xi_{ij'}|\right)^r\right]\right\}^r,$$

for some $c>0$. Then by Lemma 5 and (17)

$$\Pr(S(C_n)>\tau/3)=o\left(\frac{n^{r-\kappa r^2/\beta}}{k^r\{b(n/k)\}^{r^2}}\right)=o\left(\frac{n^{r-\kappa r^2/\beta-r^2/\alpha}}{k^{r-r^2/\alpha}}\right).$$

Let

$$\gamma=\frac{r-\frac{\kappa r^2}{\beta}-\frac{r^2}{\alpha}}{r-\frac{r^2}{\alpha}}=\frac{1-\frac{r}{\alpha}-\frac{\kappa r}{\beta}}{1-\frac{r}{\alpha}}.$$

Then, γ is smaller than 1 for all $2<\alpha<4$, as κ/β gets close to 0, and it attains its smallest value as κ/β approaches its largest possible value, i.e., $1-2/\alpha$, see (24). We now set a lower bound of γ as a function of r for α fixed,

$$\gamma_L(r;\alpha):=\frac{1-\frac{r}{\alpha}-(1-\frac{2}{\alpha})r}{1-\frac{r}{\alpha}}=\frac{\alpha-\alpha r+r}{\alpha-r}. \tag{48}$$

Since $2\beta/(\beta+2)$ in (35) is an increasing function of β and attains its upper limit when $\beta=\alpha/2$, see (24), we obtain $r<2\alpha/(\alpha+4)$. Then, since $\gamma_L(r;\alpha)$ is an decreasing function of r, γ can be arbitrarily close to $\gamma_L(2\alpha/(\alpha+4);\alpha)=(6-\alpha)/(\alpha+2)$. Thus, by Assumption 3, $k\gg n^\gamma$, and we get $\Pr(S(C_n)>\tau/3)=o(1)$. □

The following proposition is used to prove the asymptotic normality in Theorem 2. We put it in this section to help readers follow its proof easily since it uses several elements of the proof of Proposition 4. The claim is similar to (42), but $1/k$ is replaced by a suitably chosen power of k, so a more delicate argument is needed.

Proposition 5. *Suppose that Assumptions 2 3 and 4 hold. Then,*

$$d\left(\frac{1}{\sqrt{k}}\sum_{i=1}^{n} I_{\left(\widehat{R}_i/b(n/k),\ \widehat{\Theta}_i\right)},\ \frac{1}{\sqrt{k}}\sum_{i=1}^{n} I_{\left(R_i/b(n/k),\ \Theta_i\right)}\right) \xrightarrow{P} 0.$$

Proof. We follow the approach used in the proof of Proposition 4, so we skip fully analogous parts and focus on the new aspects. To get the claim, it suffices to show that

$$\Pr\left(\frac{1}{\sqrt{k}}\sum_{i=1}^{n}\left|f\left(\frac{\widehat{\mathbf{Y}}_i}{b(n/k)}\right)-f\left(\frac{\mathbf{Y}_i}{b(n/k)}\right)\right|>\tau\right)\to 0,$$

for every $f\in C_K^+([-\infty,\infty]^2\setminus\{\mathbf{0}\})$. For $0<\eta<a/2$, with a defined in (45), set

$$A_n(k):=\left\{1\le i\le n:\left\|\frac{\widehat{\mathbf{Y}}_i}{k^p b(n/k)}-\frac{\mathbf{Y}_i}{k^p b(n/k)}\right\|\le\eta,\ \left\|\frac{\mathbf{Y}_i}{k^p b(n/k)}\right\|\ge a-\eta\right\},$$

$$B_n(k):=\left\{1\le i\le n:\left\|\frac{\widehat{\mathbf{Y}}_i}{k^p b(n/k)}-\frac{\mathbf{Y}_i}{k^p b(n/k)}\right\|\le\eta,\ \left\|\frac{\mathbf{Y}_i}{k^p b(n/k)}\right\|< a-\eta\right\},$$

$$C_n(k):=\left\{1\le i\le n:\left\|\frac{\widehat{\mathbf{Y}}_i}{k^p b(n/k)}-\frac{\mathbf{Y}_i}{k^p b(n/k)}\right\|>\eta\right\},$$

where p is a positive constant such that $p\min\{r,1\}=1/2$ for some r satisfying (35). Except for the factor k^p, these sets of indexes are analogous to those used in the proof of Proposition 4. Then, we have

$$\Pr\left(\frac{1}{\sqrt{k}}\sum_{i=1}^{n}\left|f\left(\frac{\widehat{\mathbf{Y}}_i}{b(n/k)}\right)-f\left(\frac{\mathbf{Y}_i}{b(n/k)}\right)\right|>\tau\right)\le \Pr(S(A_n)>\tau/3)+\Pr(S(B_n)>\tau/3)+\Pr(S(C_n)>\tau/3),$$

where

$$S(A_n)=\frac{1}{\sqrt{k}}\sum_{i\in A_n(k)}\left|f\left(\frac{\widehat{\mathbf{Y}}_i}{b(n/k)}\right)-f\left(\frac{\mathbf{Y}_i}{b(n/k)}\right)\right|,$$

and $S(B_n)$ and $S(C_n)$ are defined analogously with $\sum_{i\in B_n(k)}$ and $\sum_{i\in C_n(k)}$, respectively. Our claim is that each of the three terms converges to 0. Before we proceed, we note some results about p in k^p to facilitate the understanding of the proofs;

$$pr=\frac{1}{2}\ \text{ for } 2<\alpha<4,\qquad p\ge\frac{1}{2}\ \text{ for } \alpha>2. \tag{49}$$

To see this, observe that $\beta<\alpha/2$ in (24) and $\frac{2\beta}{\beta+2}$ in (35) is increasing of β. It thus holds that $r<\frac{2\beta}{\beta+2}<\frac{2\alpha}{\alpha+4}$. This implies that $0<r<1$ for $2<\alpha<4$, and $0<r<2$ for $\alpha>2$.

First, observe that

$$S(A_n)\le 2\|f\|_\infty\sqrt{k}\frac{1}{k}\sum_{i=1}^{n} I_{\|\mathbf{Y}_i/k^p b(n/k)\|\ge a-\eta}=c\sqrt{k}\left(\frac{1}{k}\sum_{i=1}^{n} I_{R_i/b(n/k)\ge k^p(a-\eta)}-\nu_\alpha(k^p(a-\eta),\infty]\right)+ck^{1/2-p\alpha}(a-\eta)^{-\alpha},$$

where $\|f\|_\infty$ is defined in (46) and c is a positive constant. The last term goes to 0 since $p\alpha>p\ge 1/2$ for $\alpha>2$. Now, we focus on the first term. Assumption 4 implies

$$\mu_n:=\sqrt{k}\left(\frac{1}{k}\sum_{i=1}^{n} I_{R_i/b(n/k)}-\nu_\alpha\right)\xrightarrow{P} 0. \tag{50}$$

Consider the map g_M on $M_+(0,\infty]$, defined by $g_M(U)=U([M,\infty])$. We must show that $g_{k^p(a-\eta)}(\mu_n)\xrightarrow{P} 0$. This follows from the following more general argument. We have a sequence of signed measures on $(0,\infty]$, such that $\mu_n\xrightarrow{P} 0$. Since we can decompose μ_n into positive and negative parts, we can assume that the μ_n are positive. For $a_n\to\infty$ (in our case $a_n=k^p(a-\eta)$), we claim that $\mu_n([a_n,\infty])\xrightarrow{P} 0$. By Lemma 7, the map g_M is continuous, so for each fixed M, $\mu_n([M,\infty])\xrightarrow{P} 0$. For sufficiently large n, $a_n>1$, $\mu_n([a_n,\infty])\le\mu_n([1,\infty])$, and the claim follows.

Next, we obtain $\Pr(S(B_n)>\tau/3)=0$ in the same manner in Proposition 4.

For $S(C_n)$, we first consider the case of $\alpha > 4$. Observe that by Markov's inequality,

$$\Pr(S(C_n) > \tau/3) \le \frac{1}{\sqrt{k}} \frac{6\|f\|_\infty}{\tau\eta} \frac{n}{k^p b(n/k)} \mathrm{E}\|\widehat{\mathbf{Y}}_i - \mathbf{Y}_i\|.$$

By Lemma 5, $\mathrm{E}\|\widehat{\mathbf{Y}}_i - \mathbf{Y}_i\| = O(n^{-1/2})$, so $\Pr(S(C_n) > \tau/3) = O\left(n^{1/2-1/\alpha}/k^{1/2+p-1/\alpha}\right)$. We must thus verify that $n^{1/2-1/\alpha}/k^{1/2+p-1/\alpha} \to 0$. We know that $n^\gamma/k \to 0$ if $\gamma > \gamma(\alpha)$. We use the factorization

$$\frac{n^{1/2-1/\alpha}}{k^{1/2+p-1/\alpha}} = \left(\frac{n^\gamma}{k}\right)^{\frac{1}{2\gamma}-\frac{1}{\alpha\gamma}} \left(\frac{1}{k}\right)^{\frac{1}{2}+p-\frac{1}{\alpha}-\frac{1}{2\gamma}+\frac{1}{\alpha\gamma}}.$$

Since $\alpha > 2$, $\frac{1}{2\gamma} - \frac{1}{\alpha\gamma} > 0$, so we must be able to claim that $\frac{1}{2} + p - \frac{1}{\alpha} - \frac{1}{2\gamma} + \frac{1}{\alpha\gamma} > 0$. Since $p \ge \frac{1}{2}$, this will follow from $1 - \frac{1}{\alpha} - \frac{1}{\gamma}\left(\frac{1}{2} - \frac{1}{\alpha}\right) > 0$. A few algebraic manipulations show that the above inequality is equivalent to $\gamma > \frac{\alpha-2}{2\alpha-2} = \gamma(\alpha)$. For the case of $\alpha \in (2, 4)$, we apply Markov's and Lyapunov's inequalities, just as we did in Proposition 4. Then, by Lemma 5 and (49) we obtain

$$\Pr(S(C_n) > \tau/3) = o\left(\frac{n^{r-\kappa r^2/\beta - r^2/\alpha}}{k^{r/2+pr^2-r^2/\alpha}}\right) = o\left(\frac{n^{r-\kappa r^2/\beta - r^2/\alpha}}{k^{r-r^2/\alpha}}\right).$$

It is verified at the end of the proof of Proposition 4 that the last quantity tends to zero under Assumption 3. □

The next lemma will be used in Proposition 6 to replace $b(n/k)$ in (40) with $\widehat{R}_{(k)}$.

Lemma 10. *Under Assumptions 1, 2, and 3,* $\widehat{R}_{(k)}/b(n/k) \xrightarrow{P} 1$.

Proof. Fix $\varepsilon > 0$ and set

$$P_+(n) = \Pr\left(\frac{\widehat{R}_{(k)}}{b(n/k)} > 1+\varepsilon\right), \quad P_-(n) = \Pr\left(\frac{\widehat{R}_{(k)}}{b(n/k)} < 1-\varepsilon\right).$$

Observe that

$$P_+(n) = \Pr\left(I_{\widehat{R}_{(k)}/b(n/k)}(1+\varepsilon,\infty] = 1\right) \le \Pr\left(\frac{1}{k}\sum_{i=1}^n I_{\widehat{R}_i/b(n/k)}(1+\varepsilon,\infty] \ge 1\right) = \Pr\left(\widehat{U}_n\left((1+\varepsilon,\infty]\times\mathbb{S}^2\right) \ge 1\right).$$

A similar argument shows that $P_-(n) \le \Pr\left(\widehat{U}_n\left((1-\varepsilon,\infty]\times\mathbb{S}^2\right) < 1\right)$. The claim follows because by Lemma 7 and the continuous mapping theorem, we obtain $\widehat{U}_n((1+\varepsilon,\infty]\times\mathbb{S}^2) \xrightarrow{P} \nu_\alpha(1+\varepsilon,\infty] = (1+\varepsilon)^{-\alpha} < 1$; $\widehat{U}_n((1-\varepsilon,\infty]\times\mathbb{S}^2) \xrightarrow{P} \nu_\alpha(1-\varepsilon,\infty] = (1-\varepsilon)^{-\alpha} > 1$. □

Proposition 6. *Under Assumptions 1, 2, and 3, relation (41) holds.*

Proof. By Propositions 4 and 10, we obtain joint weak convergence $\left(\widehat{U}_n, \frac{\widehat{R}_{(k)}}{b(n/k)}\right) \Rightarrow (\nu_\alpha \times \Gamma_{jj'}, 1)$ in $M_+\left((0,\infty]\times\mathbb{S}^2\right) \times (0,\infty)$. Consider the operator $H : M_+\left((0,\infty]\times\mathbb{S}^2\right) \times (0,\infty) \to M_+((0,\infty]\times\mathbb{S}^2)$, defined by for any measurable set $A\times B \subset (0,\infty]\times\mathbb{S}^2$, $H(U,x)(A\times B) = U(xA\times B)$. Since $H\left(\widehat{U}_n, \widehat{R}_{(k)}/b(n/k)\right) = \frac{1}{k}\sum_{i=1}^n I_{\left(\widehat{R}_i/\widehat{R}_{(k)},\ \widehat{\Theta}_i\right)}$, $H\left(\nu_\alpha \times \Gamma_{jj'}, 1\right) = \nu_\alpha \times \Gamma_{jj'}$, we get (41) by Lemma 6 and the continuous mapping theorem. □

Proof of Theorem 1. Consider the map $g : M_+\left((0,\infty]\times\mathbb{S}^2\right) \to M_+\left(\mathbb{S}^2\right)$, defined by for any measurable set $A \subset \mathbb{S}^2$, $g(U) = U([1,\infty]\times A)$. Then, by Lemma 7 and the continuous mapping theorem, we obtain (27) from (41). Now we consider the map h on $M_+(\mathbb{S}^2)$ defined by $h(U) = \int_{\mathbb{S}^2} \theta_1\theta_2 U(d\boldsymbol{\theta})$. By Lemma 8 and the continuous mapping theorem, we obtain, from (27), $\int_{\mathbb{S}^2} \theta_1\theta_2 \widehat{\Gamma}_n(d\boldsymbol{\theta}) \Rightarrow \int_{\mathbb{S}^2} \theta_1\theta_2 \Gamma_{jj'}(d\boldsymbol{\theta})$. Since

$$\int_{\mathbb{S}^2} \theta_1\theta_2 \widehat{\Gamma}_n(d\boldsymbol{\theta}) = \frac{1}{k}\sum_{i=1}^n I_{R_i \ge R_{(k)}} \int_{\mathbb{S}^2} \theta_1\theta_2 I_{\Theta_i \in d\theta} = \widehat{\mathrm{D}}_n(\hat{\xi}_{ij}, \hat{\xi}_{ij'}),$$

we get the claim.

Proof of Corollary 2. Consider the map h on $M_+(\mathbb{S}^2)$ defined by

$$h(S) = \int_{\mathbb{S}^2_{(+,+)}} \theta_1\theta_2 S(d\boldsymbol{\theta}).$$

Applying the map to (27), we obtain the consistency of $\widehat{\mathrm{D}}_n^{(+,+)}(\xi_j, \xi_{j'})$ for $\mathrm{D}^{(+,+)}(\xi_j, \xi_{j'})$, by Lemma 8 and the continuous mapping theorem. The consistency of the remaining estimators can be proven in the same way, just using different quadrant domains in the map h.

Proof of Theorem 2. Define the empirical process based on the sample scores by

$$W_n(t) = \frac{1}{\sigma\sqrt{k}} \sum_{i=1}^{n} \left(\widehat{\Theta}_{i1}\widehat{\Theta}_{i2} - \mathrm{E}\left[\widetilde{\Theta}_1\widetilde{\Theta}_2\right] \right) I_{\widehat{R}_i/b(n/k) \geq t^{-1/\alpha}}, \quad t \geq 0.$$

The main argument to prove the asymptotic normality is the weak convergence of W_n to the standard Brownian motion W;

$$W_n \Rightarrow W, \quad \text{in } D[0,\infty), \tag{51}$$

where $D[0,\infty)$ is the usual Skorokhod space. Once we verify (51), then by Lemma 10 we obtain the joint convergence

$$\left(W_n(\cdot), \left(\frac{\widehat{R}_{(k)}}{b(n/k)} \right)^{-\alpha} \right) \Rightarrow (W(\cdot), 1), \quad \text{in } D[0,\infty) \times [0,\infty).$$

Applying the composition map $(x(\cdot), c) \mapsto x(c)$, we conclude that

$$\sqrt{k}\left(\widehat{\mathrm{D}}_n(\xi_j, \xi_{j'}) - \mathrm{E}\left[\widetilde{\Theta}_1\widetilde{\Theta}_2\right] \right) = \sigma W_n \left(\left(\frac{\widehat{R}_{(k)}}{b(n/k)} \right)^{-\alpha} \right) \Rightarrow \sigma W(1).$$

The general strategy is thus similar to the one employed to prove Theorem 1 in Larsson and Resnick [22]. However, in our setting, new arguments are needed to establish relations (53) and (54). These terms are zero in the proof of Larsson and Resnick [22].

Now, to show (51), consider the following decomposition

$$W_n(t) = \frac{1}{\sigma\sqrt{k}} \sum_{i=1}^{n} \left(\Theta_{i1}\Theta_{i2} - \mathrm{E}\left[\widetilde{\Theta}_1\widetilde{\Theta}_2\right] \right) I_{R_i/b(n/k) \geq t^{-1/\alpha}} + \frac{1}{\sigma\sqrt{k}} \sum_{i=1}^{n} \left(\widehat{\Theta}_{i1}\widehat{\Theta}_{i2} I_{\widehat{R}_i/b(n/k) \geq t^{-1/\alpha}} - \Theta_{i1}\Theta_{i2} I_{R_i/b(n/k) \geq t^{-1/\alpha}} \right)$$
$$+ \frac{1}{\sigma\sqrt{k}} \sum_{i=1}^{n} \mathrm{E}\left[\widetilde{\Theta}_1\widetilde{\Theta}_2\right] \left(I_{R_i/b(n/k) \geq t^{-1/\alpha}} - I_{\widehat{R}_i/b(n/k) \geq t^{-1/\alpha}} \right).$$

We will verify that

$$\frac{1}{\sigma\sqrt{k}} \sum_{i=1}^{n} \left(\Theta_{i1}\Theta_{i2} - \mathrm{E}\left[\widetilde{\Theta}_1\widetilde{\Theta}_2\right] \right) I_{R_i/b(n/k) \geq (\cdot)^{-1/\alpha}} \Rightarrow W, \quad \text{in } D[0,\infty), \tag{52}$$

and for any $s \geq 0$,

$$\sup_{0 \leq t \leq s} \left| \frac{1}{\sigma\sqrt{k}} \sum_{i=1}^{n} \left(\widehat{\Theta}_{i1}\widehat{\Theta}_{i2} I_{\widehat{R}_i/b(n/k) \geq t^{-1/\alpha}} - \Theta_{i1}\Theta_{i2} I_{R_i/b(n/k) \geq t^{-1/\alpha}} \right) \right| \xrightarrow{P} 0; \tag{53}$$

$$\mathrm{E}\left[\widetilde{\Theta}_1\widetilde{\Theta}_2\right] \sup_{0 \leq t \leq s} \left| \frac{1}{\sigma\sqrt{k}} \sum_{i=1}^{n} \left(I_{R_i/b(n/k) \geq t^{-1/\alpha}} - I_{\widehat{R}_i/b(n/k) \geq t^{-1/\alpha}} \right) \right| \xrightarrow{P} 0. \tag{54}$$

We begin with (52). Since the empirical process in (52) is based on i.i.d. population scores, if we verify

$$\sqrt{k}\left[\frac{n}{k} \Pr\left(\left(\frac{R_1}{b(n/k)}, \boldsymbol{\Theta}_1 \right) \in \cdot \right) - \frac{n}{k} \Pr\left(\frac{R_1}{b(n/k)} \in \cdot \right) \times \Gamma_{jj'} \right] \xrightarrow{v} 0, \quad \text{in } M_+((0,\infty] \times \mathbb{S}^2), \tag{55}$$

then (52) readily holds by Theorem 1 of Larsson and Resnick [22]. Their theorem is proven for nonnegative random vectors, but the proof also works for random vectors in $\mathbb{R}^d$, with a small modification.

To prove (55), we use the equivalent conditions for vague convergence presented in Theorem 3.2 of Resnick [31]. Take any relatively compact set $B \in (0,\infty]$. Then, $B \times \mathbb{S}^2$ is also relatively compact in $(0,\infty] \times \mathbb{S}^2$, so we obtain from Assumption 4,

$$\sqrt{k}\left[\frac{n}{k} \Pr\left(\frac{R_1}{b(n/k)} \in B \right) - \nu_\alpha(B) \right] \to 0. \tag{56}$$

The constant c in Assumption 4 depends on the choice of $b(n)$, so we set $c = 1$ for simplicity. Now, take any relatively compact set $A \times S \in (0,\infty] \times \mathbb{S}^2$, and observe that

$$\sqrt{k}\left[\frac{n}{k} \Pr\left(\left(\frac{R_1}{b(n/k)}, \boldsymbol{\Theta}_1 \right) \in A \times S \right) - \frac{n}{k} \Pr\left(\frac{R_1}{b(n/k)} \in A \right) \times \Gamma_{jj'}(S) \right]$$
$$= \sqrt{k}\left[\frac{n}{k} \Pr\left(\left(\frac{R_1}{b(n/k)}, \boldsymbol{\Theta}_1 \right) \in A \times S \right) - \nu_\alpha(A)\Gamma_{jj'}(S) \right] + \sqrt{k}\left[\nu_\alpha(A) - \frac{n}{k} \Pr\left(\frac{R_1}{b(n/k)} \in A \right) \right] \Gamma_{jj'}(S) \to 0.$$

The first term vanishes by Assumption 4. Also, since A is relatively compact in $(0, \infty]$ and $0 \leq \Gamma_{jj'}(S) \leq 1$, the second term goes to 0 by (56).

For (53) and (54), we Proposition 5, i.e.,

$$\frac{1}{\sqrt{k}} \sum_{i=1}^{n} I_{\left(\widehat{R}_i/b(n/k),\ \widehat{\Theta}_i\right)} - \frac{1}{\sqrt{k}} \sum_{i=1}^{n} I_{(R_i/b(n/k),\ \Theta_i)} \xrightarrow{P} 0. \tag{57}$$

Consider the map $h : M_+\left((0, \infty] \times \mathbb{S}^2\right) \to M_+(0, \infty]$, defined by $h(U) = \int_{\mathbb{S}^2} \theta_1\theta_2 U(dr, d\boldsymbol{\theta})$. Applying h to (57), by Lemma 8 and the continuous mapping theorem we obtain

$$\phi_n := \frac{1}{\sqrt{k}} \sum_{i=1}^{n} \left(\widehat{\Theta}_{i1}\widehat{\Theta}_{i2} I_{\widehat{R}_i/b(n/k)} - \Theta_{i1}\Theta_{i2} I_{R_i/b(n/k)}\right) \xrightarrow{P} 0. \tag{58}$$

We thus have a sequence of signed measures on $(0, \infty]$, such that $\phi_n \xrightarrow{P} 0$. Since a signed measure can be decomposed into positive and negative parts, we can assume that the ϕ_n are positive. Now, consider the map g_M on $M_+(0, \infty]$, defined by $g_M(U) = U([M, \infty])$. By Lemma 7, the map g_M is continuous, so for each fixed M, $\phi_n([M, \infty]) \xrightarrow{P} 0$. Therefore, for any $s \geq 0$, taking M such that $M > s$, we obtain

$$\sup_{0 \leq t \leq s} \left| \frac{1}{\sigma\sqrt{k}} \sum_{i=1}^{n} \left(\widehat{\Theta}_{i1}\widehat{\Theta}_{i2} I_{\widehat{R}_i/b(n/k) \geq t^{-1/\alpha}} - \Theta_{i1}\Theta_{i2} I_{R_i/b(n/k) \geq t^{-1/\alpha}}\right) \right| \leq \phi_n([M, \infty]) \xrightarrow{P} 0.$$

Similarly, considering the map $\ell : M_+\left((0, \infty] \times \mathbb{S}^2\right) \to M_+(0, \infty]$, defined by $\ell(U) = \int_{\mathbb{S}^2} U(dr, d\boldsymbol{\theta})$, we conclude (54).

Acknowledgments

This research was partially supported by the United States NSF grants DMS-1923142, DMS-1914882 and DMS-2123761.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jmva.2021.104887.

References

[1] G. Aneiros, R. Cao, R. Fraiman, C. Genest, P. Vieu, Recent advances in functional data analysis and high-dimensional statistics, J. Multivariate Anal. 170 (2019) 3–9.
[2] D.L. Cohn, Measure Theory, second ed., Springer, New York, 2013.
[3] D. Cooley, E. Thibaud, Decompositions of dependence for high-dimensional extremes, Biometrika 106 (2019) 587–604.
[4] S. Csörgő, P. Deheuvels, D. Mason, Kernel estimates of the tail index of a distribution, Ann. Statist. 13 (1985) 1050–1077.
[5] A. Cuevas, A partial overview of the theory of statistics with functional data, J. Statist. Plann. Inference 147 (2014) 1–23.
[6] R.A. Davis, T. Mikosch, The sample autocorrelations of heavy-tailed processes with applications to ARCH, Ann. Statist. 26 (1998) 2049–2080.
[7] R.A. Davis, T. Mikosch, The extremogram: A correlogram for extreme events, Bernoulli 15 (2009) 977–1009.
[8] R.A. Davis, S.I. Resnick, Limit theory for moving averages of random variables with regularly varying tail probabilities, Ann. Probab. 13 (1985a) 179–195.
[9] R.A. Davis, S.I. Resnick, More limit theory for the sample correlation function of moving averages, Stoch. Process. Appl. 20 (1985b) 257–279.
[10] R.A. Davis, S.I. Resnick, Limit theory for the sample covariance and correlation functions of moving averages, Ann. Statist. 14 (1986) 533–558.
[11] P. Embrechts, C. Klüppelberg, T. Mikosch, Modelling Extremal Events for Insurance and Finance, Springer, Berlin, 1997.
[12] R. de Fondeville, Functional peaks-over-threshold analysis for complex extreme events, (Ph.D. thesis), EPFL, 2018.
[13] R. de Fondeville, A.C. Davison, High-dimensional peaks-over-threshold inference, Biometrika 105 (2018) 575–592.
[14] M.G. Genton, S.A. Padoan, H. Sang, Multivariate max-stable spatial processes, Biometrika 102 (2015) 215–230.
[15] A. Goia, P. Vieu, An introduction to recent advances in high/infinite dimensional statistics, J. Multivariate Anal. 146 (2016) 1–6.
[16] E. Haeusler, J.L. Teugels, On asymptotic normality of hill's estimator for the exponent of regular variation, Ann. Statist. 13 (1985) 743–756.
[17] L. Horváth, P. Kokoszka, Inference for Functional Data with Applications, Springer, New York, 2012.
[18] T. Hsing, R. Eubank, Theoretical Foundations of Functional Data Analysis, with an Introduction To Linear Operators, Wiley, Hoboken, NJ, 2015.
[19] H. Hult, F. Lindskog, Regular variation for measures on metric spaces, in: Publications de L'Institute Mathématique. Nouvelle Série, Vol. 80, 2006, pp. 121–140.
[20] M. Kim, P. Kokoszka, Hill estimator of projections of functional data on principal components, Statistics 53 (2019) 699–720.
[21] P. Kokoszka, S. Stoev, Q. Xiong, Principal components analysis of regularly varying functions, Bernoulli 25 (2019) 3864–3882.
[22] M. Larsson, S.I. Resnick, Extremal dependence measure and extremogram: the regularly varying case, Extremes 15 (2012) 231–256.
[23] A.W. Ledford, J.A. Tawn, Statistics for near independence in multivariate extreme values, Biometrika 83 (1996) 169–187.
[24] A.W. Ledford, J.A. Tawn, Modelling dependence within joint tail regions, J. R. Stat. Soc. Ser. B Stat. Methodol. 59 (1997) 475–499.
[25] A.W. Ledford, J.A. Tawn, Diagnostics for dependence within time series extremes, J. R. Stat. Soc. Ser. B Stat. Methodol. 65 (2003) 521–543.
[26] F. Lindskog, S.I. Resnick, J. Roy, Regularly varying measures on metric spaces: Hidden regular variation and hidden jumps, Probab. Surv. 11 (2014) 270–314.
[27] T. Meiguet, Heavy tailed functional time series, (Ph.D. thesis), Universite Catholigue de Louvain, 2010.
[28] T. Opitz, Extrêmes multivariés et spatiaux: Approches spectrales et modèles elliptiques, (Ph.D. thesis), Université Montpellier 2, 2013.
[29] S.I. Resnick, Hidden regular variation, second order regular variation and asymptotic independence, Extremes 5 (2002) 303–336.
[30] S.I. Resnick, The extremal dependence measure and asymptotic independence, Stoch. Models 20 (2004) 205–227.
[31] S.I. Resnick, Heavy-Tail Phenomena, Springer, 2007.
[32] S.I. Resnick, C. Stărică, Asymptotic behavior of Hill's estimator for autoregressive data, Commun. Stat. Stoch. Models 13 (1997a) 703–721.
[33] S.I. Resnick, C. Stărică, Smoothing the hill estimator, Adv. Appl. Probab. 29 (1997b) 271–293.