Available online at www.sciencedirect.com
ScienceDirect
Fuzzy Sets and Systems 467 (2023) 108489
FUZZY sets and systems
www.elsevier.com/locate/fss

Nonparametric estimation of the multivariate Spearman's footrule: A further discussion

Ana Pérez [a,b,*], Mercedes Prieto-Alaiz [a], Fernando Chamizo [c], Eckhard Liebscher [d], Manuel Úbeda-Flores [e]

[a] *Department of Applied Economics, University of Valladolid, 47011 Valladolid, Spain*
[b] *IMUVA, University of Valladolid, Spain*
[c] *Department of Mathematics and ICMAT, Universidad Autónoma de Madrid, 28049 Madrid, Spain*
[d] *Department of Engineering and Natural Sciences, University of Applied Sciences Merseburg, D-06217 Merseburg, Germany*
[e] *Department of Mathematics, University of Almería, 04120 Almería, Spain*

Received 17 June 2022; received in revised form 21 December 2022; accepted 15 February 2023
Available online 23 February 2023

Abstract

In this paper, we propose two new estimators of the multivariate rank correlation coefficient Spearman's footrule which are based on two general estimators for Average Orthant Dependence measures. We compare the new proposals with a previous estimator existing in the literature and show that the three estimators are asymptotically equivalent, but, in small samples, one of the proposed estimators outperforms the others. We also analyse Pitman efficiency of these indices to test for multivariate independence as compared to multivariate versions of Kendall's tau and Spearman's rho.

Keywords: Spearman's footrule; Multivariate association; Concordance; Copula; Orthant dependence; Pitman efficiency

1. Introduction

Spearman's footrule is a measure of association proposed by Spearman [36] for comparing pairs of ranks, that is closely related to Spearman's rho but has been scarcely used in practice. Both coefficients can be written in terms of copulas and share some interesting properties; see Nelsen [25] for a detailed description in the bivariate case and Genest et al. [14] for a further discussion —see also the recent contributions of Genest and Jaworski [13] and Beliakov et al. [3] for bivariate copulas. However, the generalization of these coefficients to measure multivariate dependence is not straightforward since, as Durante et al. [9] point out, the pairwise properties do not always carry over to three (or more) dimensions. In this setting, and due to the increasing interest in measuring multivariate dependence, several multivariate (theoretical and empirical) versions of Spearman's rho can be found in the literature; see, for instance,

* Corresponding author at: Department of Applied Economics, University of Valladolid, C/ Avda. Valle Esgueva 6, 47011 Valladolid. Spain.
E-mail addresses: anape@uva.es (A. Pérez), prietoal@uva.es (M. Prieto-Alaiz), fernando.chamizo@uam.es (F. Chamizo), eckhard.liebscher@hs-merseburg.de (E. Liebscher), mubeda@ual.es (M. Úbeda-Flores).

https://doi.org/10.1016/j.fss.2023.02.010
0165-0114/

Nelsen ([24], [26]), Dolati and Úbeda-Flores [7], Schmid and Schmidt [32], Nelsen and Úbeda-Flores [28], García et al. [11], Pérez and Prieto-Alaiz [29], García-Gómez et al. [12] and Liebscher [21], among others. By contrast, the literature on multivariate generalizations of Spearman's footrule is scarce. Úbeda-Flores [38] proposed a copula-based multivariate version of Spearman's footrule and Genest *et al.* [14] surveyed the scattered literature on this coefficient. Pérez and Prieto-Alaiz [30] fit the multivariate Spearman's footrule into the unifying framework of the Average Orthant Dependence (AOD) measures proposed by Dolati and Úbeda-Flores [7] and prove new results. More recently, Fuchs and McCord [10] provide further results on the best lower bound of multivariate Spearman's footrule and Decancq [6] introduces two indices of diagonal multivariate dependence whose average equals the multivariate generalization of Spearman's footrule in Úbeda-Flores [38].

Interest in Spearman's footrule is partly motivated by its simplicity of calculation and its robustness, which makes it more suitable than Spearman's rho in contexts where outlying observations are likely to occur, such as those encountered in Finance, Insurance, Welfare Economics, Hydrology or Environmental Science, for example. Because of its potential use to measure multivariate dependence in these areas and the scant existing literature on its estimation, a thorough discussion on the topic is required and this is the goal of this paper. In particular, we focus on the copula-based multivariate version of Spearman's footrule proposed by Úbeda-Flores [38] and further develop the problem of its estimation, beyond the succinct formula provided in that paper.[1] To do so, we exploit that Spearman's footrule belongs to the class of AOD measures and resort to some results for this class in Dolati and Úbeda-Flores [7].

Our contribution is fourfold. First, we propose two new nonparametric estimators which are based on two sample versions of the AOD measures proposed by Dolati and Úbeda-Flores [7], and provide new insights on the estimator in Úbeda-Flores [38], which was the only one existing in previous literature. We prove that, in the bivariate and trivariate cases, one of our estimators coincides with that in Úbeda-Flores [38], but this coincidence does not longer hold for more than three dimensions. Moreover, we show that, under independence, this new estimator is unbiased regardless of the dimensions considered. Second, we demonstrate that the three estimators at hand are asymptotically equivalent, and we prove its asymptotic normality under milder conditions than those in Genest et al. [14], using a central limit theorem for non-degenerate U-statistics rather than the functional approach used by these authors. Third, we illustrate that the three estimators differ on their finite sample properties, through extensive Monte Carlo experiments based on some parametric well-known copulas. Our results show that, as expected, both bias and dispersion reduce as the sample size increases and the performance of the three estimators is quite similar in large samples. However, in small samples, one of the two estimators we propose clearly outperforms the other, especially in terms of bias, and it is nearly equivalent to the estimator already proposed in Úbeda-Flores [38]. Finally, we explore the possibility of using the estimators of the Spearman's footrule as tests statistics for multivariate independence and carry out asymptotic relative efficiency comparisons with other well-known measures of association, like Spearman's rho and Kendall's tau, under local copula-based alternatives. Although no general recommendation can be made, our results are promising.

The rest of the paper is organized as follows. Section 2 briefly reviews the copula function and introduces the Spearman's footrule as an AOD measure of multivariate association and summarizes its main properties. Section 3 is devoted to nonparametric estimation of Spearman's footrule. We introduce two new estimators based on sample versions of the AOD measures and discuss their properties as compared to the other estimator previously proposed in the literature. Section 4 includes Monte Carlo experiments to compare finite sample performance of the estimators discussed in Section 3. In Section 5 the proposed estimators are also investigated in the light of their Pitman asymptotic relative efficiency, as compared to Spearman's rho and Kendall's tau, when they are used as statistics for testing multivariate independence. Finally, Section 6 gathers a summary of our main conclusions.

2. Preliminary concepts

For $d \geq 2$, we consider the d-dimensional random vector $\mathbf{X} = (X_1, \ldots, X_d)$ with joint distribution function H, univariate marginal distribution functions $F_1, \ldots, F_d$, and copula C. Sklar's Theorem (see Sklar [35]) says that H can be represented as

$$H(\mathbf{x}) = C(F_1(x_1), \ldots, F_d(x_d)) \quad \text{for } \mathbf{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d, \tag{1}$$

[1] Mesfioui and Quessy [23] propose a non-continuous multivariate version of Spearman's footrule, but do not tackle its estimation, as they recognized that this is a difficult problem, even in the continuous case.

where the copula function $C : [0,1]^d \to [0,1]$ is a d-dimensional joint distribution function whose margins are standard uniform $U(0,1)$. Throughout this paper, we will assume that the margins $F_1, \dots, F_d$ are all continuous and so the copula C in (1) is unique. For a complete survey on copulas, we refer to the monographs [8,27].

Let us define the probability integral transformations, $U_j = F_j(X_j)$, with $j = 1, \dots, d$, which are uniform $U(0,1)$ and whose joint distribution function is the copula C, i.e., let $\mathbf{U} = (U_1, \dots, U_d) \sim C$. Two examples of copulas are the independent copula Π —or product copula— and the comonotonic copula M. The former denotes the copula of d independent random variables and is defined as $\Pi(\mathbf{u}) = \prod_{j=1}^{d} u_j$, for any real vector $\mathbf{u} = (u_1, \dots, u_d) \in [0,1]^d$. The copula M represents maximal dependence, i.e., the case when each of the random variables $X_1, \dots, X_d$ is almost surely a strictly increasing function of any of the others. This copula and its associated survival function, $\overline{M}$, are defined, for any real vector $\mathbf{u} = (u_1, \dots, u_d) \in [0,1]^d$, as

$$M(\mathbf{u}) = \min_{1 \le j \le d} \{u_j\}, \tag{2}$$

$$\overline{M}(\mathbf{u}) = 1 - \max_{1 \le j \le d} \{u_j\}. \tag{3}$$

Let $Q'_d(C, M)$ be the probability of concordance between C and M defined as (see Nelsen [26])

$$Q'_d(C, M) = \int\limits_{[0,1]^d} \left(M(\mathbf{u}) + \overline{M}(\mathbf{u}) \right) \, \mathrm{d}C(\mathbf{u}).$$

For any d-copula C, the multivariate Spearman's footrule $\varphi_d(C)$ —or simply φ_d, if there is no confusion—, proposed by Úbeda-Flores [38] can be defined as

$$\varphi_d = \frac{Q'_d(C, M) - a_d}{b_d - a_d}, \tag{4}$$

where

$$a_d = Q'_d(\Pi, M) = \int_{[0,1]^d} \left(M(\mathbf{u}) + \overline{M}(\mathbf{u}) \right) \, \mathrm{d}\Pi(\mathbf{u}) = \frac{2}{d+1}$$

and

$$b_d = Q'_d(M, M) = \int_0^1 \left(M(t, \dots, t) + \overline{M}(t, \dots, t) \right) \, \mathrm{d}t = 1;$$

that is,

$$\varphi_d = 1 - \frac{(d+1)(1 - Q'_d(C, M))}{(d-1)}. \tag{5}$$

Notice that φ_d can be alternatively written as

$$\varphi_d = 1 - \frac{d+1}{d-1} \int\limits_{[0,1]^d} \left(\max_{1 \le j \le d} \{u_j\} - \min_{1 \le j \le d} \{u_j\} \right) \, \mathrm{d}C(\mathbf{u}). \tag{6}$$

The coefficient φ_d in (4) is a particular member of the general class of AOD measures of multivariate association introduced by Dolati and Úbeda-Flores [7]. In particular, φ_d can be regarded as a normalized probability of concordance between the distribution of $\mathbf{X}$, as represented by their copula C, and the copula M, which represents maximal dependence. When $C = \Pi$ (the case of independent variables), we have $\varphi_d = 0$ and when $C = M$ (maximal dependence), we have $\varphi_d = 1$, and the inequality $\varphi_d \ge -1/d$ always holds. In the bivariate case ($d = 2$), if the underlying copula is the lower Fréchet-Hoeffding W, Spearman's footrule attains its best-possible lower bound, that is, $\varphi_2 = -1/2$. For higher dimensions ($d \ge 3$), Fuchs and McCord [10] provides characterizations of the copulas attaining the best lower bound $-1/d$ of multivariate Spearman's footrule. Moreover, in the bivariate case ($d = 2$), the coefficient φ_d in (5) becomes the bivariate footrule,

$$\varphi_2 = 6 \int\limits_{[0,1]^2} M(u_1, u_2) \, \mathrm{d}C(u_1, u_2) - 2,$$

whereas in the trivariate case ($d = 3$), the coefficient φ_3 can be written as

$$\varphi_3 = \frac{\varphi_2^{1,2} + \varphi_2^{1,3} + \varphi_2^{2,3}}{3}, \tag{7}$$

where $\varphi_2^{i,j}$ denotes the corresponding Spearman's footrule for the bivariate random variable (X_i, X_j), with $1 \le i < j \le 3$.

The following examples illustrate the values of φ_d for some d-copulas.

Example 1. For $d \ge 2$, let $C_\theta^{\mathbf{FGM}}$ be the d-copula given by

$$C_\theta^{\mathbf{FGM}}(\mathbf{u}) = \left(\prod_{i=1}^{d} u_i\right)\left[1 + \theta \prod_{i=1}^{d}(1 - u_i)\right], \quad \mathbf{u} \in [0,1]^d, \tag{8}$$

with θ in $[0, 1]$. $C_\theta^{\mathbf{FGM}}$ belongs to the *Farlie-Gumbel-Morgenstern* family of d-copulas (see Durante and Sempi [8] and Nelsen [27] for more details). We note that all the margins of any dimension $j \ge 2$ of (8) are Π, in the corresponding dimension j. Then we have

$$\varphi_d\left(C_\theta^{\mathbf{FGM}}\right) = \theta \frac{\left(1 + (-1)^d\right)(d+1)(d!)^2}{(d-1)(2d+1)!}; \tag{9}$$

in particular,

$$\varphi_2\left(C_\theta^{\mathbf{FGM}}\right) = \frac{\theta}{5}, \quad \varphi_3\left(C_\theta^{\mathbf{FGM}}\right) = 0, \quad \varphi_4\left(C_\theta^{\mathbf{FGM}}\right) = \frac{\theta}{189}, \ldots$$

Example 2. For $d \ge 2$, let $C_\theta^{\mathbf{CA}}$ be a multivariate generalization of the (parametric) Cuadras-Augé family of copulas (see Cuadras and Augé [5]) given by

$$C_\theta^{\mathbf{CA}}(\mathbf{u}) = (\Pi(\mathbf{u}))^{1-\theta}(M(\mathbf{u}))^{\theta}$$

for all $\mathbf{u} \in [0,1]^d$ and any $\theta \in [0,1]$, for which $C_0^{\mathbf{CA}} = \Pi$ and $C_1^{\mathbf{CA}} = M$. For this d-copula, all the margins of any dimension $j \ge 2$ belong again to the corresponding Cuadras-Augé family of j-copulas. Thus, after some algebra, we obtain

$$\varphi_d\left(C_\theta^{\mathbf{CA}}\right) = 1 + \frac{d+1}{d-1}\left(\frac{1}{d+1-(d-1)\theta} - \frac{1}{1+\theta} + \frac{d!}{(1-\theta)\prod_{i=0}^{d}(\beta+i)}\right),$$

where $\beta = (1+\theta)/(1-\theta)$, for $\theta \in [0, 1[$. Note that, for instance,

$$\varphi_2\left(C_\theta^{\mathbf{CA}}\right) = \varphi_3\left(C_\theta^{\mathbf{CA}}\right) = \frac{2\theta}{3-\theta}, \quad \varphi_4\left(C_\theta^{\mathbf{CA}}\right) = -\frac{\theta(21\theta^2 - 70\theta + 61)}{9\theta^3 - 60\theta^2 + 129\theta - 90}, \ldots$$

Example 3. For this example, we need to recall the concept of Archimedean d-copula. Let ϕ be a continuous strictly decreasing function from $[0, 1]$ to $[0, \infty]$ such that $\phi(0) = \infty$ and $\phi(1) = 0$, and let ϕ^{-1} be the inverse of ϕ. Then the function given by

$$C_\phi(\mathbf{u}) = \phi^{-1}\left(\sum_{i=1}^{d} \phi(u_i)\right), \quad \mathbf{u} \in [0,1]^d,$$

is a d-copula if, and only if, ϕ^{-1} is completely monotonic on $[0, \infty[$, i.e., $(-1)^k \frac{d^k}{dt^k}\phi^{-1}(t) \ge 0$ for all $t \in]0, \infty[$, $k = 1, 2, \ldots$, where $\frac{d^k}{dt^k}$ denotes the k-th derivative. In such a case, we say that C_ϕ is an *Archimedean d-copula*, and the function ϕ is called a *generator* of C_ϕ. For more details, see McNeil and Nešlehová [22] and Nelsen [27]. Then, for $d \ge 2$, we have

$$\varphi_d\left(C_\phi\right) = \frac{-d^2+d-2}{2(d-1)} + \frac{d+1}{d-1}\left[\int_0^1 \phi^{-1}\left(d\,\phi(t)\right)\,\mathrm{d}t + \sum_{i=2}^{d}(-1)^i\binom{d}{i}\int_0^1 \phi^{-1}\left(i\,\phi(t)\right)\,\mathrm{d}t\right].$$

Next we show the cases of two particular Archimedean d-copulas.

1. Consider the generator $\phi_\theta(t) = t^{-\theta} - 1$ for all $t \geq 0$ and $\theta > 0$, which generates a subfamily of the bivariate Clayton family of copulas (see Clayton [4]). Since $\phi_\theta^{-1}(t) = (1+t)^{-1/\theta}$ and

$$(-1)^k \frac{d^k}{dt^k}\phi_\theta^{-1}(t) = \frac{(1+t)^{-k-1/\theta}}{\theta^k}\prod_{i=1}^{k-1}(1+k\theta) \geq 0$$

for all $t \geq 0$, then ϕ_θ^{-1} is completely monotonic on $[0, +\infty[$, and thus we obtain the d-copula

$$C_\theta^{\mathbf{C}}(\mathbf{u}) = \left(\sum_{i=1}^{d} u_i^{-\theta} - d + 1\right)^{-1/\theta}, \quad \mathbf{u} \in [0,1]^d, \tag{10}$$

where $C_\theta^{\mathbf{C}}(\mathbf{u}) = 0$ when $u_i = 0$ for some $i \in \{1, \ldots, d\}$. This d-copula is a generalization of the Clayton family of 2-copulas for which $\lim_{\theta\to 0^+} C_\theta^{\mathbf{C}}(\mathbf{u}) = \Pi(\mathbf{u})$ and $C_\theta^{\mathbf{C}}(\mathbf{u}) \geq \Pi(\mathbf{u})$ for all $\mathbf{u} \in [0,1]^d$ (see [27] for more details). Therefore, after some tedious —but elementary— algebra, we have

$$\begin{aligned}\varphi_d\left(C_\theta^{\mathbf{C}}\right) = {} & \frac{-d^2+d-2}{2(d-1)} \\ & + \frac{d+1}{2(d-1)}\left[d^{-1/\theta}\,{}_2F_1\left(\frac{1}{\theta}, \frac{2}{\theta}; \frac{2}{\theta}+1; \frac{d-1}{d}\right)\right. \\ & \left. + \sum_{i=2}^{d}(-1)^i\binom{d}{i} i^{-1/\theta}\,{}_2F_1\left(\frac{1}{\theta}, \frac{2}{\theta}; \frac{2}{\theta}+1; \frac{i-1}{i}\right)\right],\end{aligned}$$

where ${}_2F_1$ denotes the (Gaussian) hypergeometric function (see, e.g., Seaborn [33]). Since all the bivariate margins of (10) —which we denote by $C_{ij,\theta}^{\mathbf{C}}$, for $1 \leq i < j \leq d$— are equal, i.e.,

$$C_{ij,\theta}^{\mathbf{C}}(u,v) = \left(u^{-\theta} + v^{-\theta} - 1\right)^{-1/\theta} \tag{11}$$

for all $(u,v) \in [0,1]^2$, we have

$$\varphi_2\left(C_\theta^{\mathbf{C}}\right) = \varphi_3\left(C_\theta^{\mathbf{C}}\right) = 3 \cdot 2^{-1/\theta}\,{}_2F_1\left(\frac{1}{\theta}, \frac{2}{\theta}; \frac{2}{\theta}+1; \frac{1}{2}\right) - 2$$

and

$$\begin{aligned}\varphi_4\left(C_\theta^{\mathbf{C}}\right) = {} & -\frac{7}{3} + 5 \cdot 2^{-1/\theta}\,{}_2F_1\left(\frac{1}{\theta}, \frac{2}{\theta}; \frac{2}{\theta}+1; \frac{1}{2}\right) - \frac{10}{3} \cdot 3^{-1/\theta}\,{}_2F_1\left(\frac{1}{\theta}, \frac{2}{\theta}; \frac{2}{\theta}+1; \frac{2}{3}\right) \\ & + \frac{5}{3} \cdot 4^{-1/\theta}\,{}_2F_1\left(\frac{1}{\theta}, \frac{2}{\theta}; \frac{2}{\theta}+1; \frac{3}{4}\right).\end{aligned}$$

2. Consider the generator $\phi_\theta(t) = \ln\left(\frac{1-\theta(1-t)}{t}\right)$ for all $t \geq 0$, with $\theta \in [0,1[$. Since $\phi_\theta^{-1}(t) = \frac{1-\theta}{e^t-\theta}$ and

$$(-1)^k \frac{d^k}{dt^k}\phi_\theta^{-1}(t) = \frac{(1-\theta)e^t \sum_{i=0}^{k-1} p_i \theta^i e^{(k-i-1)t}}{(e^t-\theta)^{k+1}} \geq 0$$

for all $t \in]0, +\infty[$, where $p_i > 0$ for all $i = 0, 1, \ldots, k-1$, with $p_0 = p_{k-1} = 1$ (see Lemma 1 in Appendix A), then ϕ_θ^{-1} is completely monotonic for $\theta \in [0,1[$ and

$$C_\theta^{\mathbf{AMH}}(\mathbf{u}) = \phi_\theta^{-1}\left(\sum_{i=1}^{d}\phi_\theta(u_i)\right) = (1-\theta)\left[\prod_{i=1}^{d}\left(\frac{1-\theta}{u_i}+\theta\right)-\theta\right]^{-1}, \quad \mathbf{u}\in[0,1]^d, \tag{12}$$

is a d-copula (note that $C_0^{\mathbf{AMH}} = \Pi$), which is a generalization of the *Ali-Mikhail-Haq (AMH, for short) family of bivariate copulas* given by

$$C_{2,\theta}^{\mathbf{AMH}}(u,v) = \frac{uv}{1-\theta(1-u)(1-v)} \tag{13}$$

for all $(u,v)\in[0,1]^2$ (see Ali et al. [1] and Nelsen [27]). We note that all the margins of any dimension $j \geq 2$ of (12) belong to the corresponding AMH family of j-copulas. Therefore,

$$\varphi_d\left(C_\theta^{\mathbf{AMH}}\right) = \frac{-d^2+d-2}{2(d-1)} + \frac{(d+1)(1-\theta)}{d-1}\left[\int_0^1 \frac{t^d}{(\theta t-\theta+1)^d-\theta t^d}\,\mathrm{d}t + \sum_{i=2}^{d}(-1)^i\binom{d}{i}\int_0^1 \frac{t^i}{(\theta t-\theta+1)^i-\theta t^i}\,\mathrm{d}t\right]; \tag{14}$$

and since all the bivariate margins of (12) are the 2-copulas given by (13), then we have

$$\varphi_2\left(C_\theta^{\mathbf{AMH}}\right) = \varphi_3\left(C_\theta^{\mathbf{AMH}}\right) = 6\int_0^1 \frac{t^2}{1-\theta(1-t)^2}\,\mathrm{d}t - 2 = 6\left(\frac{1+\theta}{2\theta^{3/2}}\ln\frac{1+\sqrt{\theta}}{1-\sqrt{\theta}} + \frac{1}{\theta}\ln(1-\theta) - \frac{1}{\theta}\right) - 2$$

(see Gradshteyn and Ryzhik [17, 2.175.4]) and

$$\varphi_4\left(C_\theta^{\mathbf{AMH}}\right) = \frac{10(1-\theta)}{3}\int_0^1 \sum_{i=2}^{4}\frac{(-1)^i(5-i)t^i}{(\theta t-\theta+1)^i-\theta t^i}\,\mathrm{d}t - \frac{7}{3}.$$

Fig. 1 depicts the behaviour of the Spearman's footrule coefficient, $\varphi_d(C_\theta)$, as a function of θ for the d-copulas studied throughout this section, with $d = \{3,4,6\}$. Notice that the y-axis scale for the FGM copula is different, as this copula models small departures from independence and so it results in much smaller values of the Spearman's footrule than the others. The y-axis scale for the AMH is also different as this copula does not allow for higher dependencies either.

3. Estimation

In practice, the copula C is unknown and φ_d must be estimated from the data. Therefore, a sample version of φ_d is required. In order to do that, let $\mathbf{X}_i = (X_{i1},\dots,X_{id})$, $i = 1,\dots,n$, be a sample of n serially independent random vectors from the d-dimensional vector $\mathbf{X} = (X_1,\dots,X_d)$, and let $(R_{i1},\dots,R_{id})$, $i = 1,\dots,n$, denote the associated vectors of componentwise ranks of the sample, such that R_{ij} is the *rank* of X_{ij} among $\{X_{1j},\dots,X_{nj}\}$, with $j = 1,\dots,d$. Then, the d-variate empirical copula, C_n, is defined as the empirical cumulative distribution function computed from the scaled ranks, i.e.,

$$C_n(\mathbf{u}) = \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}\left\{\frac{R_{i1}}{n+1}\leq u_1,\dots,\frac{R_{id}}{n+1}\leq u_d\right\},$$

for all $\mathbf{u}\in[0,1]^d$, where $\mathbf{1}\{A\}$ denotes the indicator function of a set A.

In this section, we first review the definition and main properties of the only estimator of φ_d previously proposed in the literature (to our knowledge), that will be called $\widehat{\varphi}_{nd}^{(1)}$. Then, given that φ_d is an AOD measure, we draw on the two sample versions for AOD measures proposed in Dolati and Úbeda-Flores [7] and apply them to build up naturally two new estimators of Spearman's footrule, say $\widehat{\varphi}_{nd}^{(2)}$ and $\widehat{\varphi}_{nd}^{(3)}$. The first one consists of only estimating the function

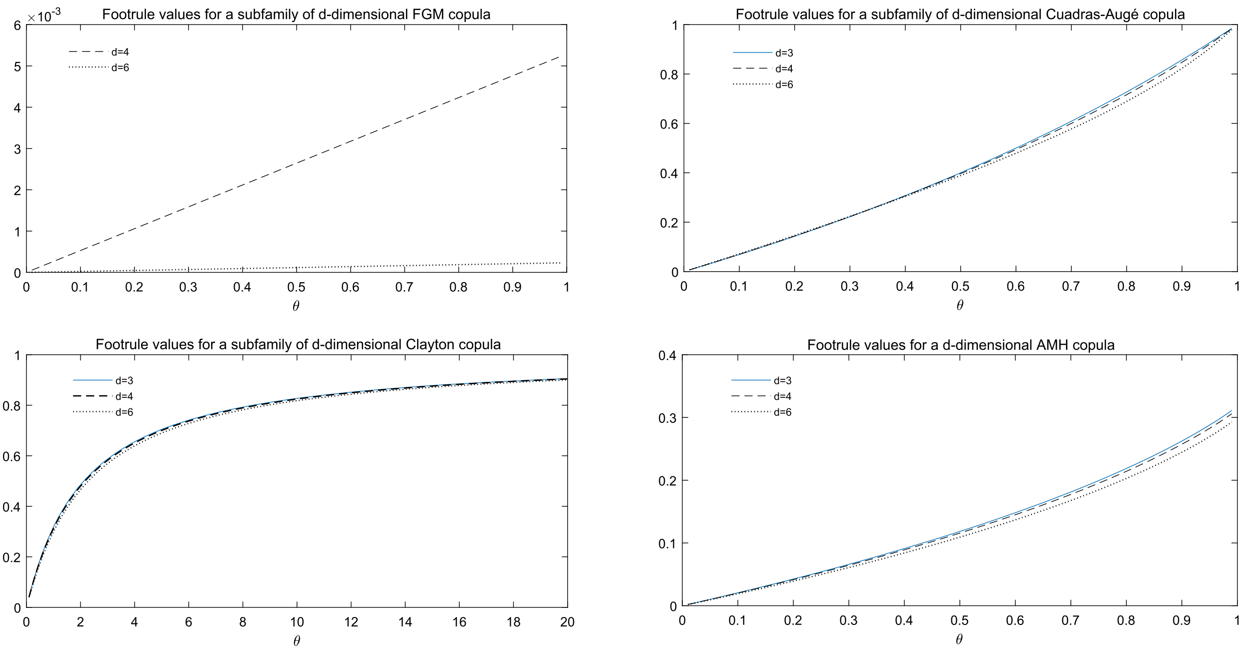

Fig. 1. Behaviour of the d-dimensional Spearman's footrule for three different dimensions, $d = \{3, 4, 6\}$, and four copulas: FMG, Cuadras-Augé, Clayton and AMH.

$Q'_d(C, M)$ in (4), whereas the second one is obtained by estimating in (4) the function $Q'_d(C, M)$ as well as the coefficients a_d and b_d. We show that the three estimators at hand are asymptotically equivalent and so the asymptotic normality of only one of them, namely $\widehat{\varphi}^{(2)}_{nd}$, is proved.

3.1. First estimator

As far as we know, Úbeda-Flores [38] proposed the only empirical version of φ_d, given by

$$\widehat{\varphi}^{(1)}_{nd} = 1 - \frac{d+1}{d-1}\frac{1}{n^2-1}\sum_{i=1}^{n} L_i, \tag{15}$$

where, for each $i \in \{1, ..., n\}$, L_i is defined as

$$L_i = \max_{1\le j\le d}(R_{ij}) - \min_{1\le j\le d}(R_{ij}). \tag{16}$$

As expected, when the ranks in each dimension coincide, i.e., in the case of perfect positive dependence, $\widehat{\varphi}^{(1)}_{nd} = 1$. Due to the lack of motivation in the original proposal of this estimator, we will discuss later two other estimators based on expression (4). Before that, we briefly summarize the main properties of $\widehat{\varphi}^{(1)}_{nd}$.

3.1.1. Particular cases

- When $d = 2$, the expression in (15) reduces to the sample bivariate Spearman's footrule —known as f_S— given by

$$\widehat{\varphi}^{(1)}_{n2} := f_S = 1 - \frac{3}{n^2-1}\sum_{i=1}^{n}|R_{i1} - R_{i2}|. \tag{17}$$

- In the trivariate case ($d = 3$), the sample version of Spearman's footrule in (15) is equal to the average of the three pairwise sample Spearman's footrule coefficients, that is, property (7) continues to hold for the corresponding empirical coefficients; see Pérez and Prieto-Alaiz [30].

3.1.2. Asymptotic distribution

Genest *et al.* [14] consider an alternative expression of the estimator $\widehat{\varphi}_{nd}^{(1)}$, which is based on (6), namely

$$\widehat{\varphi}_{nd}^{(1)} = 1 - \frac{d+1}{d-1}\frac{n}{n-1} \int_{[0,1]^d} \left(\max_{1\le j\le d} \{u_j\} - \min_{1\le j\le d}\{u_j\} \right) dC_n(\mathbf{u}),$$

and show that, under fairly general conditions —the d-copula C admits continuous (first-order) partial derivatives on $]0,1[^d$—, $\widehat{\varphi}_{nd}^{(1)}$ is asymptotically unbiased, though it is biased in finite samples. Moreover, they show that $\widehat{\varphi}_{nd}^{(1)}$ is asymptotically normally distributed, i.e.,

$$\sqrt{n}\left(\widehat{\varphi}_{nd}^{(1)} - \varphi_d\right) \xrightarrow{d} \mathcal{N}\left(0, \sigma^2_{\varphi_C}\right), \tag{18}$$

where $\xrightarrow{d}$ denotes the convergence in distribution and $\sigma^2_{\varphi_C}$ is defined in equation (A5) in Genest *et al.* [14].

3.1.3. Moments at independence

Genest *et al.* [14] show that, under the hypothesis of independence, that is, when the underlying copula is Π, we have

$$E_\Pi\left(\widehat{\varphi}_{nd}^{(1)}\right) = 1 - \frac{d+1}{d-1}\frac{n}{n-1}\left\{1 - \frac{2}{(n+1)n^d}\sum_{i=1}^{n} i^d\right\}. \tag{19}$$

Noticeably, when $d = \{2,3\}$, this expectation vanishes, i.e., $E_\Pi\left(\widehat{\varphi}_{n2}^{(1)}\right) = E_\Pi\left(\widehat{\varphi}_{n3}^{(1)}\right) = 0$. When $d > 3$, this expectation is only $O(1/n^2)$. For instance, when $d = 4$, it becomes $E_\Pi\left(\widehat{\varphi}_{n4}^{(1)}\right) = 1/(9n^2)$. Hence, under independence, the estimator $\widehat{\varphi}_{nd}^{(1)}$ is unbiased for $d = \{2,3\}$ but it is biased for $d > 3$. Genest *et al.* [14] also provides the expression of the large-sample variance of the asymptotic distribution in (18) when the underlying copula is Π, which is

$$\sigma_\Pi^2 = 2\left(\frac{d+1}{d-1}\right)^2\left\{\frac{2+4d-d^2+d^3}{d(d+2)(2d+1)(d+1)^2} - \frac{\mathcal{B}(d,d+2)}{d+1}\right\},$$

where $\mathcal{B}$ denotes the Beta function. This can be rearranged as

$$\sigma_\Pi^2 = \frac{2}{d(2d+1)}\left(\frac{d+1}{d-1}\right)^2\left(1 - \frac{d(5d+1)}{(d+2)(d+1)^2} - \binom{2d}{d}^{-1}\right), \tag{20}$$

giving $\sigma_\Pi^2 = \{2/5, 2/15, 149/2268, 11/280\}$ for $d = \{2,3,4,5\}$, respectively, and showing that σ_Π^2 decays as $1/d^2$ when d grows.

3.2. Second estimator

To look for alternative empirical versions of φ_d, we will make use of the fact that φ_d is an AOD measure and we will propose two new estimators based on the two sample versions for AOD measures given in Dolati and Úbeda-Flores [7]. In doing so, our goal is twofold. First, we want to compare the properties of the new proposals with the estimator $\widehat{\varphi}_{nd}^{(1)}$ previously discussed, but we also intend to shed light on the definition of $\widehat{\varphi}_{nd}^{(1)}$.

One possible estimator of φ_d based on (4) is given by

$$\widehat{\varphi}_{nd}^{(2)} = \frac{\widehat{Q}'_{nd}(C,M) - a_d}{b_d - a_d}, \tag{21}$$

where the estimator $\widehat{Q}'_{nd}$ for Q'_d is given by

$$\begin{aligned}\widehat{Q}'_{nd}(C,M) &= \frac{1}{n}\sum_{i=1}^{n}\left(M\left(\frac{R_{i1}}{n+1},\ldots,\frac{R_{id}}{n+1}\right) + \overline{M}\left(\frac{R_{i1}}{n+1},\ldots,\frac{R_{id}}{n+1}\right)\right)\\ &= 1 - \frac{1}{n(n+1)}\sum_{i=1}^{n} L_i,\end{aligned} \tag{22}$$

where L_i is defined in (16). Hence, $\widehat{\varphi}_{nd}^{(2)}$ can be written as

$$\widehat{\varphi}_{nd}^{(2)} = 1 - \frac{d+1}{d-1}\frac{1}{n(n+1)}\sum_{i=1}^{n} L_i. \tag{23}$$

This estimator is closely related to the estimator $\widehat{\varphi}_{nd}^{(1)}$ previously discussed. Actually, the following relationship holds:

$$\widehat{\varphi}_{nd}^{(2)} = \frac{1}{n} + \left(1 - \frac{1}{n}\right)\widehat{\varphi}_{nd}^{(1)}. \tag{24}$$

Hence, both estimators are asymptotically equivalent. Moreover, when the ranks in each dimension coincide, i.e., in the case of perfect positive dependence, $\widehat{\varphi}_{nd}^{(1)} = \widehat{\varphi}_{nd}^{(2)} = 1$, otherwise, $\widehat{\varphi}_{nd}^{(2)} > \widehat{\varphi}_{nd}^{(1)}$. Therefore, if $\widehat{\varphi}_{nd}^{(1)}$ attains its best-possible lower bound, $\widehat{\varphi}_{nd}^{(1)} = -1/d$, then $\widehat{\varphi}_{nd}^{(2)}$ will not, since the inequality $\widehat{\varphi}_{nd}^{(2)} > -1/d$ always holds. Notice also that, in the bivariate case, $\widehat{\varphi}_{n2}^{(1)}$ attains the best-possible lower bound, $\widehat{\varphi}_{n2}^{(1)} = -1/2$, in case the ranks in one dimension are just the reversed of the other dimension and with n odd, but $\widehat{\varphi}_{n2}^{(2)}$ does not.

3.2.1. Particular cases

- When $d = 2$, the coefficient in (23) becomes

$$\widehat{\varphi}_{n2}^{(2)} = 1 - \frac{3}{n(n+1)}\sum_{i=1}^{n}|R_{i1} - R_{i2}| = 1 - \frac{n-1}{n}\frac{3}{(n^2-1)}\sum_{i=1}^{n}|R_{i1} - R_{i2}|,$$

  which does not exactly coincide with the sample bivariate Spearman's footrule in (17) due to the (negligible) factor $(n-1)/n$.
- When $d = 3$, the coefficient in (23) is not equal to the average of the three pairwise Spearman's footrule coefficients, that is, in this case, property (7) does not hold for the corresponding empirical coefficients.

3.2.2. Asymptotic distribution

In order to derive the asymptotic distribution of $\widehat{\varphi}_{nd}^{(2)}$, we first notice that, from (4) and (21), we have:

$$\sqrt{n}\left(\widehat{\varphi}_{nd}^{(2)} - \varphi_d\right) = \frac{d+1}{d-1}\sqrt{n}\left(\widehat{Q}'_{nd}(C,M) - Q'_d(C,M)\right).$$

We introduce the following assumption:

Assumption $\mathcal{A}_C$. Let C_{jk} be the bivariate copula of the random pair $\left(X_j, X_k\right)$, for $1 \le j < k \le d$. Assume that the partial derivatives of C_{jk} with respect to the first argument —denoted by $C_{jk}^{(1)}$— exist and are continuous on a neighbourhood $\left\{(u,v) \in [0,1]^2 : |u-v| < \eta\right\}$ of the diagonal, where $\eta > 0$. □

Remark 1. Let $C_\theta^{\mathbf{AMH}}$ be the generalization of the AMH d-copula given by (12). Since all the bivariate margins are given by (13), then we have

$$C_{jk,\theta}^{\mathbf{AMH}\,(1)}(u,v) = \frac{v(1-\theta(1-v))}{(1-\theta(1-u)(1-v))^2}.$$

Observe that the domain is the set given by $\left\{(u,v) \in [0,1]^2 : \theta(1-u)(1-v) \ne 1\right\}$, and the bivariate margins of $C_\theta^{\mathbf{AMH}}$ fulfil Assumption $\mathcal{A}_C$. Also note that the bivariate margins of the FGM family of d-copulas given in Example 1 trivially satisfy Assumption $\mathcal{A}_C$.

However, the generalization of the Cuadras-Augé d-copula given in Example 2 does not, since it has a singular component (coming from M) such that $C_{jk}^{\mathbf{CA}\,(1)}$ does not exist at the diagonal (see Nelsen [27] for details). Furthermore, the generalization of the Clayton d-copulas given by (10) does not satisfy the assumption either since, from (11), we have

$$C_{jk,\theta}^{\mathbf{C}\,(1)}(u,v) = u^{-\theta-1}\left(u^{-\theta} + v^{-\theta} - 1\right)^{-1/\theta-1},$$

so that $l := \lim_{(u,v)\to(0,0)} C^{\mathbf{C}\,(1)}_{jk,\theta}(u,v)$ is different for different trajectories, e.g. $l = 2^{-(\theta+1)/\theta}$ for $v = u$ and $l = 3^{-(\theta+1)/\theta}$ for $v = 2u$; therefore $C^{\mathbf{C}\,(1)}_{jk,\theta}$ is not continuous at the origin.

The following theorem provides the asymptotic normality for $\widehat{Q}'_{nd}(C, M)$ and $\widehat{\varphi}^{(2)}_{nd}$.

Theorem 1. *Assume that $F_1, F_2, \ldots, F_d$ are continuous and, for any d-copula C, $\mathcal{A}_C$ is fulfilled. Then*

$$\sqrt{n}\left(\widehat{Q}'_{nd}(C, M) - Q'_d(C, M)\right) \xrightarrow{d} \mathcal{N}\left(0, \sigma^2\right),$$

where σ^2 is given by expression (50) in Appendix B. Moreover,

$$\sqrt{n}\left(\widehat{\varphi}^{(2)}_{nd} - \varphi_d\right) \xrightarrow{d} \mathcal{N}\left(0, \left(\frac{d+1}{d-1}\right)^2 \sigma^2\right). \tag{25}$$

Proof. See Appendix B. □

As expected, the asymptotic distribution in (25) coincides with the asymptotic distribution of $\widehat{\varphi}^{(1)}_{nd}$ derived in Genest et al. [14], but Theorem 1 above requires milder conditions. Proposition 3 of Genest et al. [14] uses the stronger assumption that the partial derivatives of the whole copula $C : [0, 1]^d \to [0, 1]$ are continuous. Moreover, Genest et al. [14, Proposition 3] provide a sophisticated expression for the asymptotic variance of $\widehat{\varphi}^{(1)}_{nd}$ depending on covariances of Gaussian processes defined in terms of the copula. By contrast, our formula $\left(\frac{d+1}{d-1}\right)^2 \sigma^2$ for the asymptotic variance of $\widehat{\varphi}^{(2)}_{nd}$ in (25) is explicit through the expression (50) for σ^2. Furthermore, the variance σ^2 in Theorem 1 above can be estimated as explained in Appendix C.

In the paper by Genest et al. [14] the functional approach is applied for proving asymptotic normality. In the proof of our Theorem 1, the crucial point is the use of a central limit theorem for non-degenerate U-statistics. This central limit theorem is proved utilizing Hoeffding's projection method.

Theorem 1 can be used to establish confidence intervals for φ_d. Moreover, tests about φ_d can be constructed, for example the test of the hypothesis $H_0 : \varphi_d \geq K$, $H_1 : \varphi_d < K$, where $K > 0$ is a given number.

3.2.3. Moments at independence

Taking into account the relationship in (24) and the formula of $E_\Pi\left(\widehat{\varphi}^{(1)}_{nd}\right)$ in (19), it turns out that, under independence, the estimator $\widehat{\varphi}^{(2)}_{nd}$ is always biased, even for $d = \{2, 3\}$. In particular, we have that, under independence

$$E_\Pi\left(\widehat{\varphi}^{(2)}_{n2}\right) = E(\widehat{\varphi}^{(2)}_{n3}) = \frac{1}{n},$$
$$E_\Pi\left(\widehat{\varphi}^{(2)}_{n4}\right) = \frac{9n^2 + n - 1}{9n^3}.$$

To sum up, the estimator $\widehat{\varphi}^{(2)}_{nd}$ is well-motivated and it is asymptotically equivalent to the estimator $\widehat{\varphi}^{(1)}_{nd}$, but the latter has better properties in finite samples. Moreover, the former takes a narrower range of values than it should be.

3.3. Third estimator

The last estimator of the coefficient φ_d we consider, is based on estimating in (4) both the function $Q'_d(C, M)$ and the coefficients a_d and b_d, that is,

$$\widehat{\varphi}^{(3)}_{nd} = \frac{\widehat{Q}'_{nd}(C, M) - \widehat{a}_{nd}}{\widehat{b}_{nd} - \widehat{a}_{nd}}, \tag{26}$$

where $\widehat{Q}'_{nd}(C, M)$ is given in (22) and

$$\widehat{a}_{nd} = \frac{1}{n^d} \sum_{i_1=1}^{n} \cdots \sum_{i_d=1}^{n} \left[M\left(\frac{i_1}{n+1}, \ldots, \frac{i_d}{n+1}\right) + \overline{M}\left(\frac{i_1}{n+1}, \ldots, \frac{i_d}{n+1}\right) \right],$$
$$\widehat{b}_{nd} = \frac{1}{n} \sum_{i=1}^{n} \left[M\left(\frac{i}{n+1}, \ldots, \frac{i}{n+1}\right) + \overline{M}\left(\frac{i}{n+1}, \ldots, \frac{i}{n+1}\right) \right].$$

Now, from (2) and (3), we obtain:

$$\widehat{a}_{nd} = 1 - \frac{1}{n^d(n+1)} \sum_{i_1=1}^{n} \cdots \sum_{i_d=1}^{n} [\max(i_1, \ldots, i_d) - \min(i_1, \ldots, i_d)],$$
$$\widehat{b}_{nd} = 1,$$

and, putting these expressions back into (26), we have:

$$\widehat{\varphi}_{nd}^{(3)} = 1 - \frac{n^{d-1}}{S_{nd}} \sum_{i=1}^{n} L_i, \tag{27}$$

where L_i is defined in (16) and S_{nd} is as follows (see Appendix D):

$$S_{nd} = (n+1)n^d - 2\sum_{i=1}^{n} i^d = \frac{d-1}{d+1} n^{d+1} - \frac{2}{d+1} \sum_{k=1}^{\lfloor d/2 \rfloor} \binom{d+1}{2k} B_{2k} n^{d+1-2k}, \tag{28}$$

where $\lfloor x \rfloor$ denotes the floor of x and B_m is the corresponding Bernoulli number, i.e., $\{B_{2k}\}_{k=1}^{\infty} = \{1/6, -1/30, 1/42, -1/30, 5/66, \ldots\}$.

Now, the question arises on how the estimator $\widehat{\varphi}_{nd}^{(1)}$ in (15) compares with the new estimator $\widehat{\varphi}_{nd}^{(3)}$ in (27). As expected, when the ranks in each dimension coincide, i.e., in the case of perfect dependence, both estimators coincide and attain their maximum value, i.e., $\widehat{\varphi}_{nd}^{(1)} = \widehat{\varphi}_{nd}^{(3)} = 1$. Moreover, the following relationship holds:

$$\widehat{\varphi}_{nd}^{(3)} = 1 + \frac{d-1}{d+1} \frac{(n^2-1)n^{d-1}}{S_{nd}} (\widehat{\varphi}_{nd}^{(1)} - 1).$$

3.3.1. Particular cases

- In the bivariate case ($d = 2$), the expression in (28) becomes $S_{n2} = n(n^2-1)/3$. Now, putting this back into (27), we obtain

  $$\widehat{\varphi}_{n2}^{(3)} = 1 - \frac{n}{S_{n2}} \sum_{i=1}^{n} |R_{i1} - R_{i2}| = 1 - \frac{3}{(n^2-1)} \sum_{i=1}^{n} |R_{i1} - R_{i2}|,$$

  which is the sample bivariate Spearman's footrule f_S in (17). Therefore, when $d = 2$, both the estimator in (15) and that in (27) coincide and reduce to the sample bivariate Spearman's footrule in (17), i.e. $\widehat{\varphi}_{n2}^{(1)} = \widehat{\varphi}_{n2}^{(3)} = f_S$.
- In the trivariate case ($d = 3$), the expression in (28) becomes $S_{n3} = n^2(n^2-1)/2$, and so,

  $$\widehat{\varphi}_{n3}^{(3)} = 1 - \frac{n^2}{S_{n3}} \sum_{i=1}^{n} L_i = 1 - \frac{2}{(n^2-1)} \sum_{i=1}^{n} L_i.$$

  Noticeably, if we evaluate the estimator $\widehat{\varphi}_{nd}^{(1)}$ in (15) for $d = 3$, the expression above also comes up. That is, when $d = 3$, both the coefficient in (15) and that in (27) coincide, i.e. $\widehat{\varphi}_{n3}^{(1)} = \widehat{\varphi}_{n3}^{(3)}$, and both are equal to the average of the three pairwise sample Spearman's footrule coefficients.
- In higher dimensions ($d > 3$), both estimators no longer coincide unless they are both equals to 1. The proof is as follows. Using (15) and (27), it turns out that $\widehat{\varphi}_{nd}^{(1)} = \widehat{\varphi}_{nd}^{(3)}$ if and only if

  $$(n^2-1) = \frac{d+1}{d-1} \frac{S_{nd}}{n^{d-1}}. \tag{29}$$

But, taking into account (28), the left hand side of (29) becomes

$$\frac{d+1}{d-1}\frac{S_{nd}}{n^{d-1}} = n^2 - c_d + \epsilon(n,d),$$

where

$$c_d = \frac{2B_2}{d-1}\binom{d+1}{2} = \frac{d(d+1)}{6(d-1)},$$

and $\epsilon(n,d)$ is given, for $d > 3$, by

$$\epsilon(n,d) = \frac{2}{d-1}\sum_{k=2}^{\lfloor d/2 \rfloor}\binom{d+1}{2k}B_{2k}n^{2-2k} = O(n^{-2}).$$

Therefore, for $d > 3$, the estimators $\widehat{\varphi}_{nd}^{(1)}$ and $\widehat{\varphi}_{nd}^{(3)}$ will not coincide unless $\epsilon(n,d) = 0$ and $c_d = 1$, but the former is not possible and the latter implies $d(d+1) = 6(d-1)$, and the roots of this quadratic equation are $d = \{2, 3\}$.

3.3.2. Asymptotic distribution

Notice that, as $n^2 - 1$ and $n^2 - c_d + \epsilon(n,d)$ are asymptotically equal (meaning that their quotient tends to 1 as $n \to \infty$), the estimators $\widehat{\varphi}_{nd}^{(1)}$ and $\widehat{\varphi}_{nd}^{(3)}$ are asymptotically equivalent. Even more, we have $n\big(\widehat{\varphi}_{nd}^{(1)} - \widehat{\varphi}_{nd}^{(3)}\big) \to 0$ because

$$\limsup_{n\to\infty} n\big|\widehat{\varphi}_{nd}^{(1)} - \widehat{\varphi}_{nd}^{(3)}\big| = \frac{d+1}{d-1}\limsup_{n\to\infty}\left(\left|\frac{n}{n^2-1} - \frac{n}{n^2-c_d}\right|\sum_{i=1}^{n} L_i\right)$$

and $0 \le L_i \le n$ implies that the limit above is zero. A similar argument shows that, in fact, $n^{\alpha}\big(\widehat{\varphi}_{nd}^{(1)} - \widehat{\varphi}_{nd}^{(3)}\big) \to 0$ for any $\alpha < 2$.

Remark 2. We want to note that, since the random variables are measurable functions, the limits are at almost every point, whence the convergence is almost surely (a.s.).

3.3.3. Moments at independence

From the result on the expectation of $\widehat{\varphi}_{nd}^{(1)}$ in (19), it turns out that, under independence,

$$E_{\Pi}\left(\sum_{i=1}^{n} L_i\right) = n(n+1) - \frac{2}{n^{d-1}}\sum_{i=1}^{n} i^d.$$

Therefore, taking into account (27), when the underlying copula is Π, we have

$$E_{\Pi}\left(\widehat{\varphi}_{nd}^{(3)}\right) = 1 - \frac{n^{d-1}}{S_{nd}}\left[n(n+1) - \frac{2}{n^{d-1}}\sum_{i=1}^{n} i^d\right],$$

and the first equality for S_{nd} in (28) shows that the expression above vanishes. That is, under independence, the estimator $\widehat{\varphi}_{nd}^{(3)}$, unlike the two estimators previously discussed, is always unbiased, regardless of the value of d, i.e., we have

$$E_{\Pi}\left(\widehat{\varphi}_{nd}^{(3)}\right) = 0.$$

To summarize, the new estimators $\widehat{\varphi}_{nd}^{(2)}$ and $\widehat{\varphi}_{nd}^{(3)}$ are well-motivated and they are asymptotically equivalent to the estimator $\widehat{\varphi}_{nd}^{(1)}$ first introduced by Úbeda-Flores [38]. Moreover, $\widehat{\varphi}_{nd}^{(1)}$ and $\widehat{\varphi}_{nd}^{(3)}$ share some interesting properties. For instance, they coincide for $d = \{2, 3\}$. Actually, in the bivariate case ($d = 2$), they both reduce to the well-known empirical bivariate Spearman's footrule whereas in the trivariate case ($d = 3$), they both become the average of the corresponding pairwise estimators. By contrast, the estimator $\widehat{\varphi}_{nd}^{(2)}$ does not fulfil any of these properties. Noticeably, the estimator $\widehat{\varphi}_{nd}^{(3)}$ has generally better finite sample properties. For instance, under the hypothesis of independence, it is always unbiased, regardless of the dimension d, whereas the others are not: $\widehat{\varphi}_{nd}^{(2)}$ is always biased and $\widehat{\varphi}_{nd}^{(1)}$ is only unbiased for $d = \{2, 3\}$, in which case it coincides with $\widehat{\varphi}_{nd}^{(3)}$. Hence, it seems that, among the three estimators considered, $\widehat{\varphi}_{nd}^{(3)}$ is preferable, especially in small samples, but $\widehat{\varphi}_{nd}^{(1)}$ also makes sense and can be a useful estimator in the bivariate and trivariate case, where it coincides with $\widehat{\varphi}_{nd}^{(3)}$.

4. Finite sample comparisons

This section is devoted to present the results of a simulation study of the three estimators discussed in the previous sections in order to evaluate their practical performance. The results presented below are quite suggestive of the applicability of the large-sample theory to finite samples.

We carry out a set of Monte Carlo simulations for the FGM d-copula described in Example 1 and the AMH d-copula described in Example 3 —the Monte Carlo simulations have been implemented using the R package copula in Hofert et al. [18]. As a benchmark, we also simulate the d-dimensional independent copula Π. We consider three different dimensions, $d = \{3, 4, 5\}$, and five sample sizes, $n = \{10, 20, 50, 100, 500\}$. For each copula, each parameter value and each dimension d, we generate 1000 Monte Carlo replicates of size n and for each replicate, we compute the three estimators of the coefficient φ_d defined in the previous section, namely $\widehat{\varphi}_{nd}^{(1)}$, $\widehat{\varphi}_{nd}^{(2)}$ and $\widehat{\varphi}_{nd}^{(3)}$. For the sake of simplicity, we only fully report the results for $d = 4$ and $n = \{50, 500\}$. The results for $d = 3$ and $d = 5$ and other sample sizes are not displayed here to save space, but they are available as additional material. Interestingly, conclusions remain unchanged.

Tables 1 and 2 report summary statistics for the three estimators under all considered models with $n = 50$ and $n = 500$, respectively. These tables display, for each copula and estimator, the true multivariate Spearman's footrule (true), the mean, the standard deviation (sd), the root mean square error (rmse), the 2.5^{th} percentile ($q_{2.5}$), the first quartile (q_{25}), the median (q_{50}), the third quartile (q_{75}) and the 97.5^{th} percentile ($q_{97.5}$). The investigated copulas are the independence copula, which appears under the case $\theta = 0$, the FGM copula with parameter values $\theta = \{0.25, 0.5, 0.75, 0.95\}$ and the AMH copula with parameter values $\theta = \{0.2, 0.4, 0.6, 0.8\}$. To enhance the global picture of the finite performance of the three estimators, Fig. 2 and Fig. 3 display the box-plots of the empirical distribution of the three estimators for FGM and AMH copulas, respectively, with dimension $d = 4$ and three sample sizes $n = \{50, 100, 500\}$. The true value of the multivariate Spearman's footrule for each simulated model is also displayed as a horizontal line. The independence copula will be further analyzed in Fig. 4.

As we can see, when $n = 50$, the small sample performance of $\widehat{\varphi}_{nd}^{(2)}$ is quite different from the other two estimators, which display hardly no differences between them. First, $\widehat{\varphi}_{nd}^{(1)}$ and $\widehat{\varphi}_{nd}^{(3)}$ tend to underestimate the true footrule, whereas $\widehat{\varphi}_{nd}^{(2)}$ tends to overestimate it. Second, there is a clear superior performance of the former estimators due to their smaller bias. For example, in the AMH model, the maximum relative bias does not exceed 7% in both $\widehat{\varphi}_{nd}^{(1)}$ and $\widehat{\varphi}_{nd}^{(3)}$ while the relative bias in $\widehat{\varphi}_{nd}^{(2)}$ could reach nearly 40%. The median follows the same pattern as the sample mean for the three estimators, which suggests a symmetric sampling distribution, a feature confirmed in Figs. 2 and 3. So, the larger bias of $\widehat{\varphi}_{nd}^{(2)}$ is not due to asymmetry of its sampling distribution, but it could be related to the relationship between the three estimators set in Section 3. At a glance, the more limited values that $\widehat{\varphi}_{nd}^{(2)}$ can take —recall that $\widehat{\varphi}_{nd}^{(1)} < \widehat{\varphi}_{nd}^{(2)}$— push its sampling distribution up, making this estimator not so well behaved, in terms of bias, as the other two estimators.

By contrast, in terms of dispersion, the differences between the three estimators are not so marked. The standard deviation of three estimators reported in Table 1 are nearly the same, so their performance in terms of rmse is similar to that in terms of bias. Moreover, the length of the intervals with the 50% and 95% of values of the three estimators are very similar, although the quantiles which defined such intervals are quite different in the case of $\widehat{\varphi}_{nd}^{(2)}$.

The large sample performance of the three estimators drawn from the Monte Carlo simulations is consistent with our theoretical results; see Table 2 and the box-plots shown in Fig. 2 and Fig. 3. As expected, as the sample size increases, bias and dispersion tend to reduce and the differences between the three estimators become negligible. The same conclusion emerges comparing the quantiles: looking across the entries of the Table 2, we can observe that the quantiles of the three estimators are pretty equal for all models considered, although the sampling distribution of $\widehat{\varphi}_{nd}^{(2)}$ seems to be still slightly upwards.

Finally, to get a better insight on the approximation to the asymptotic distribution, Fig. 4 displays the empirical distribution of the three estimators (rescaled by $\sqrt{n}$) for the independent copula Π and $d = 4$, with the corresponding

Table 1
Bias, standard deviation, rmse and different quantiles of three estimators of the Spearman's footrule coefficient. Computations are based on 1000 Monte Carlo replicates of size $n = 50$ of the Farlie–Gumbel–Morgenstern d-copula with $d = 4$ and parameter values $\theta = \{0.25, 0.5, 0.75, 1\}$ and the Ali-Mikhail-Haq d-copula with $d = 4$ and parameter values $\theta = \{0.2, 0.4, 0.6, 0.8\}$. The case $\theta = 0$ stands for independence.

Independence copula										
θ	**true**	**estimator**	**mean**	**sd**	**rmse**	**$q_{2.5}$**	**q_{25}**	**q_{50}**	**q_{75}**	**$q_{97.5}$**
0	**0**	$\widehat{\varphi_1}$	-0.0001	0.0345	0.0345	-0.0664	-0.0232	-0.0004	0.0216	0.0676
		$\widehat{\varphi_2}$	0.0199	0.0338	0.0392	-0.0451	-0.0028	0.0196	0.0412	0.0863
		$\widehat{\varphi_3}$	-0.0002	0.0345	0.0345	-0.0665	-0.0233	-0.0004	0.0216	0.0676
Farlie–Gumbel–Morgenstern copula										
θ	**true**	**estimator**	**mean**	**sd**	**rmse**	**$q_{2.5}$**	**q_{25}**	**q_{50}**	**q_{75}**	**$q_{97.5}$**
0.25	**0.0013**	$\widehat{\varphi_1}$	0.0013	0.0352	0.0352	-0.0671	-0.0231	0.0009	0.0249	0.0532
		$\widehat{\varphi_2}$	0.0213	0.0345	0.0399	-0.0458	-0.0026	0.0209	0.0444	0.0627
		$\widehat{\varphi_3}$	0.0013	0.0352	0.0352	-0.0672	-0.0231	0.0009	0.0249	0.0532
0.5	**0.0026**	$\widehat{\varphi_1}$	0.0038	0.0364	0.0364	-0.0638	-0.0211	0.0026	0.0283	0.0543
		$\widehat{\varphi_2}$	0.0238	0.0357	0.0415	-0.0425	-0.0007	0.0225	0.0477	0.0637
		$\widehat{\varphi_3}$	0.0038	0.0364	0.0364	-0.0638	-0.0211	0.0026	0.0282	0.0542
0.75	**0.004**	$\widehat{\varphi_1}$	0.0027	0.0368	0.0368	-0.0671	-0.0224	0.0016	0.0276	0.0574
		$\widehat{\varphi_2}$	0.0226	0.036	0.0406	-0.0458	-0.002	0.0216	0.0471	0.0668
		$\widehat{\varphi_3}$	0.0026	0.0368	0.0368	-0.0671	-0.0225	0.0016	0.0276	0.0574
0.95	**0.005**	$\widehat{\varphi_1}$	0.0052	0.0384	0.0384	-0.0671	-0.0204	0.0036	0.0316	0.0587
		$\widehat{\varphi_2}$	0.0251	0.0377	0.0427	-0.0458	0	0.0235	0.051	0.0682
		$\widehat{\varphi_3}$	0.0051	0.0384	0.0384	-0.0672	-0.0205	0.0036	0.0316	0.0587
Ali-Mikhail-Haq copula										
θ	**true**	**estimator**	**mean**	**sd**	**rmse**	**$q_{2.5}$**	**q_{25}**	**q_{50}**	**q_{75}**	**$q_{97.5}$**
0.2	**0.0416**	$\widehat{\varphi_1}$	0.0389	0.0424	0.0425	-0.0318	0.0123	0.0383	0.0651	0.1243
		$\widehat{\varphi_2}$	0.0581	0.0415	0.0447	-0.0111	0.032	0.0575	0.0838	0.1418
		$\widehat{\varphi_3}$	0.0388	0.0424	0.0425	-0.0318	0.0122	0.0382	0.0651	0.1243
0.4	**0.089**	$\widehat{\varphi_1}$	0.0849	0.045	0.0451	0.0029	0.0548	0.0876	0.1183	0.1797
		$\widehat{\varphi_2}$	0.1032	0.0441	0.0463	0.0228	0.0737	0.1059	0.1359	0.1961
		$\widehat{\varphi_3}$	0.0849	0.045	0.0451	0.0029	0.0547	0.0876	0.1183	0.1797
0.6	**0.1449**	$\widehat{\varphi_1}$	0.1382	0.0491	0.0495	0.053	0.1076	0.141	0.1725	0.239
		$\widehat{\varphi_2}$	0.1554	0.0481	0.0492	0.0719	0.1255	0.1582	0.1891	0.2542
		$\widehat{\varphi_3}$	0.1382	0.0491	0.0495	0.0529	0.1076	0.141	0.1725	0.239
0.8	**0.2142**	$\widehat{\varphi_1}$	0.2093	0.0496	0.0498	0.1129	0.1755	0.2104	0.2444	0.3077
		$\widehat{\varphi_2}$	0.2251	0.0486	0.0498	0.1306	0.192	0.2261	0.2595	0.3216
		$\widehat{\varphi_3}$	0.2092	0.0496	0.0498	0.1128	0.1755	0.2103	0.2443	0.3077

asymptotic Normal distribution. Note that closed-form expressions for the asymptotic variance are rare, but in the case of independence, we do have an explicit formula for the asymptotic variance; see equation (20). Noticeably, the approximation to the Normal distribution for the estimators $\widehat{\varphi}_{nd}^{(1)}$ and $\widehat{\varphi}_{nd}^{(3)}$ is very good, even for small samples. However, that is not the case for the estimator $\widehat{\varphi}_{nd}^{(2)}$, which is clearly upward biased, as previously remarked. Nevertheless, as the sample size increases, the bias decreases, the differences between the three estimators reduce and the Normal approximation becomes appropriate for the three of them.

To sump up, the results do suggest that, in small samples, $\widehat{\varphi}_{nd}^{(1)}$ and $\widehat{\varphi}_{nd}^{(3)}$ outperform $\widehat{\varphi}_{nd}^{(2)}$ since the former are systematically less biased and the dispersion of the three estimators are similar. However, in large samples, the differences between the three estimators are quite negligible and, so, large sample theory can be applied to evaluate their practical performance.

Table 2
Bias, standard deviation, rmse and different quantiles of three estimators of the Spearman's footrule coefficient. Computations are based on 1000 Monte Carlo replicates of size $n = 500$ of the Ali-Mikhai-Haq d-copula with $d = 4$ and parameter values $\theta = \{0.2, 0.4, 0.6, 0.8\}$ and the Farlie–Gumbel–Morgenstern d-copula with $d = 4$ and parameter values $\theta = \{0.25, 0.5, 0.75, 1\}$. The case $\theta = 0$ stands for independence.

Independence copula									
θ	**true**	**estimator**	**mean**	**sd**	**rmse**	$q_{2.5}$	q_{25}	q_{50}	q_{75}	$q_{97.5}$
0	**0**	$\widehat{\varphi_1}$	-0.0004	0.0112	0.0112	-0.0223	-0.0079	-0.0007	0.0073	0.0215
		$\widehat{\varphi_2}$	0.0016	0.0112	0.0113	-0.0202	-0.0059	0.0013	0.0093	0.0234
		$\widehat{\varphi_3}$	-0.0004	0.0112	0.0112	-0.0223	-0.0079	-0.0007	0.0073	0.0214
Farlie–Gumbel–Morgenstern copula										
θ	**true**	**estimator**	**mean**	**sd**	**rmse**	$q_{2.5}$	q_{25}	q_{50}	q_{75}	$q_{97.5}$
0.25	**0.0013**	$\widehat{\varphi_1}$	0.0002	0.0117	0.0117	-0.023	-0.0073	0.0002	0.0081	0.0251
		$\widehat{\varphi_2}$	0.0022	0.0116	0.0117	-0.021	-0.0053	0.0022	0.0101	0.027
		$\widehat{\varphi_3}$	0.0002	0.0117	0.0117	-0.023	-0.0073	0.0002	0.0081	0.0251
0.5	**0.0026**	$\widehat{\varphi_1}$	0.0026	0.0119	0.0119	-0.0199	-0.0057	0.0027	0.0111	0.0256
		$\widehat{\varphi_2}$	0.0046	0.0119	0.012	-0.0178	-0.0036	0.0047	0.0131	0.0275
		$\widehat{\varphi_3}$	0.0026	0.0119	0.0119	-0.0199	-0.0057	0.0027	0.0111	0.0256
0.75	**0.004**	$\widehat{\varphi_1}$	0.0039	0.0117	0.0117	-0.0171	-0.0046	0.0035	0.0115	0.0281
		$\widehat{\varphi_2}$	0.0058	0.0116	0.0118	-0.015	-0.0026	0.0055	0.0135	0.03
		$\widehat{\varphi_3}$	0.0039	0.0117	0.0117	-0.0171	-0.0046	0.0035	0.0115	0.0281
0.95	**0.005**	$\widehat{\varphi_1}$	0.0049	0.0117	0.0117	-0.0165	-0.0034	0.0045	0.0128	0.0284
		$\widehat{\varphi_2}$	0.0068	0.0117	0.0118	-0.0145	-0.0014	0.0065	0.0148	0.0303
		$\widehat{\varphi_3}$	0.0049	0.0117	0.0117	-0.0165	-0.0034	0.0045	0.0128	0.0284
Ali-Mikhail-Haq copula										
θ	**true**	**estimator**	**mean**	**sd**	**rmse**	$q_{2.5}$	q_{25}	q_{50}	q_{75}	$q_{97.5}$
0.2	**0.0416**	$\widehat{\varphi_1}$	0.0413	0.0127	0.0127	0.0162	0.0324	0.0412	0.0497	0.0648
		$\widehat{\varphi_2}$	0.0432	0.0127	0.0128	0.0182	0.0343	0.0431	0.0516	0.0667
		$\widehat{\varphi_3}$	0.0413	0.0127	0.0127	0.0162	0.0324	0.0412	0.0497	0.0648
0.4	**0.089**	$\widehat{\varphi_1}$	0.0893	0.0137	0.0137	0.0618	0.0799	0.089	0.0983	0.1158
		$\widehat{\varphi_2}$	0.0911	0.0136	0.0138	0.0636	0.0817	0.0908	0.1001	0.1175
		$\widehat{\varphi_3}$	0.0893	0.0137	0.0137	0.0618	0.0799	0.089	0.0983	0.1158
0.6	**0.1449**	$\widehat{\varphi_1}$	0.1440	0.0146	0.0147	0.1167	0.1339	0.1439	0.1557	0.1736
		$\widehat{\varphi_2}$	0.1457	0.0146	0.0146	0.1185	0.1356	0.1456	0.1574	0.1752
		$\widehat{\varphi_3}$	0.1440	0.0146	0.0147	0.1167	0.1339	0.1439	0.1557	0.1736
0.8	**0.2142**	$\widehat{\varphi_1}$	0.2134	0.0158	0.0158	0.1839	0.2038	0.2141	0.224	0.2458
		$\widehat{\varphi_2}$	0.2150	0.0157	0.0158	0.1855	0.2054	0.2157	0.2255	0.2473
		$\widehat{\varphi_3}$	0.2134	0.0158	0.0158	0.1839	0.2038	0.2141	0.224	0.2458

5. Asymptotic relative efficiency

Empirical Spearman's footrule, as well as other copula-based measures, like Spearman's rho or Kendall's tau, are natural statistics for testing independence. Hence, it makes sense to compare them to help the user make a substantiated choice of the most efficient one. In this section, we undertake such comparison by looking at their Pitman's asymptotic relative efficiencies (ARE) as test statistics for independence. In particular, we will compute Pitman's ARE using the classical formulation for asymptotically normally distributed statistics; see Serfling [34]. This requires introducing some further notation. First, let C_θ denote a one-parameter d-copula such that $C_0 = \Pi$, so the hypothesis of independence can be regarded as $H_0 : \theta = 0$, and let $\varphi_d = \varphi_d(\theta)$ be the population Spearman's footrule for this

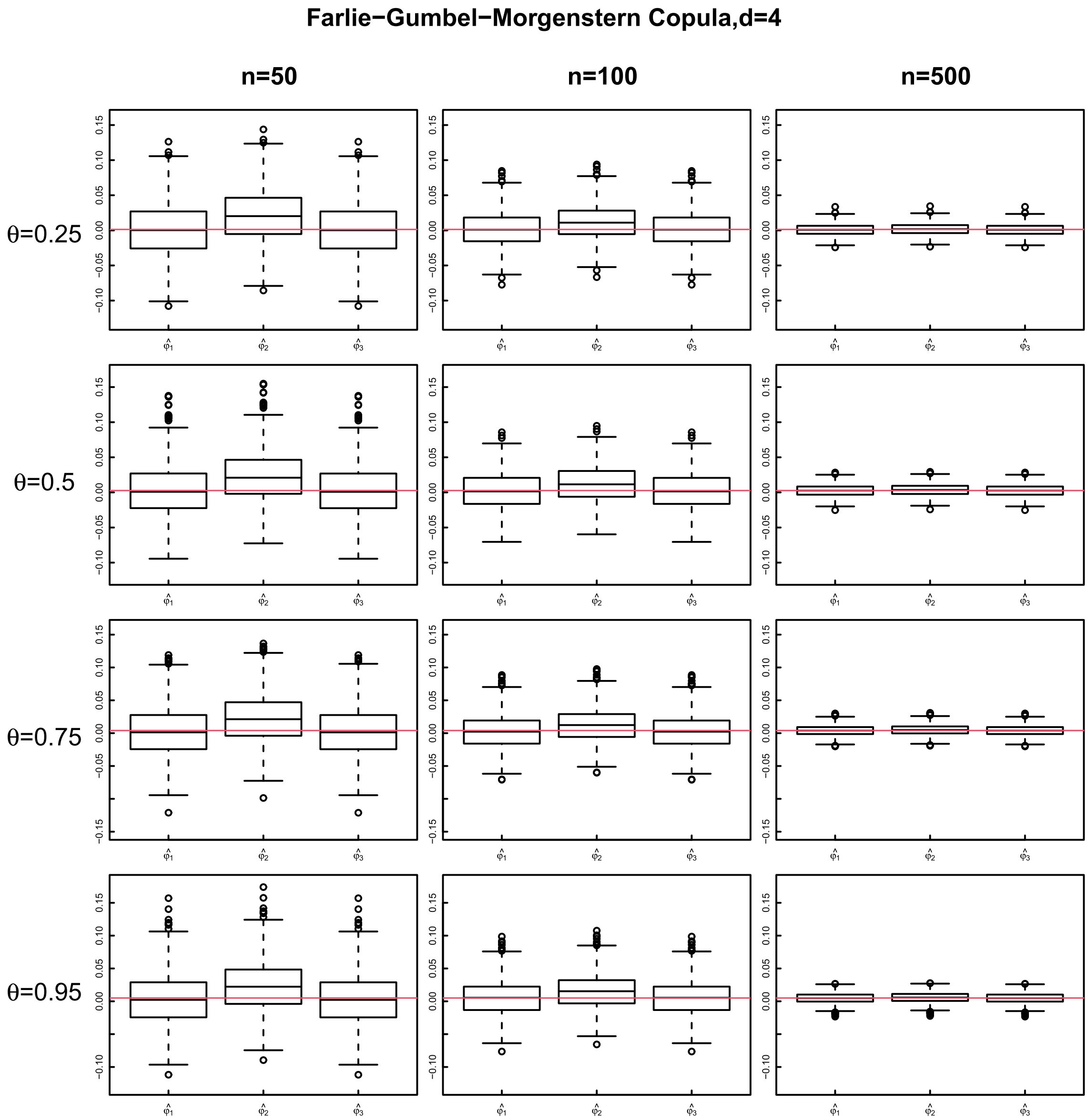

Fig. 2. Empirical distribution of three estimators of Spearman's footrule coefficient for the Farlie–Gumbel–Morgenstern d-copula with $d = 4$, parameter values $\theta = \{0.25, 0.5, 0.75, 0.95\}$ and sample sizes $n = \{50, 100, 500\}$. Horizontal red line in each graph sets the true value of Spearman's footrule.

copula C_θ. Second, given that the three empirical versions of Spearman's footrule in Section 3 are asymptotically equivalent, here ownwards we will drop the superscript and denote the empirical footrule as $\widehat{\varphi}_{nd}$. From the results in Section 3, we know that $\sqrt{n}\,(\widehat{\varphi}_{nd} - \varphi_d)$ is an asymptotically zero-mean normal variable, whose limiting variance, say $\sigma_\varphi^2(\theta)$, depends on the underlying copula C_θ. Let V_n be another statistic for independence, such that $\sqrt{n}\,(V_n - \mu_V(\theta))$ is also an asymptotically zero-mean normal variable with limiting variance $\sigma_V^2(\theta)$ under C_θ. Thus, provided that the copula C_θ meets mild regularity conditions, the Pitman's ARE of $\widehat{\varphi}_{nd}$ compared with V_n will be calculated as

$$ARE(\widehat{\varphi}_{nd}, V_n) = \left(\frac{\varphi_d'(0)}{\sigma_\varphi(0)}\right)^2 \Bigg/ \left(\frac{\mu_V'(0)}{\sigma_V(0)}\right)^2, \tag{30}$$

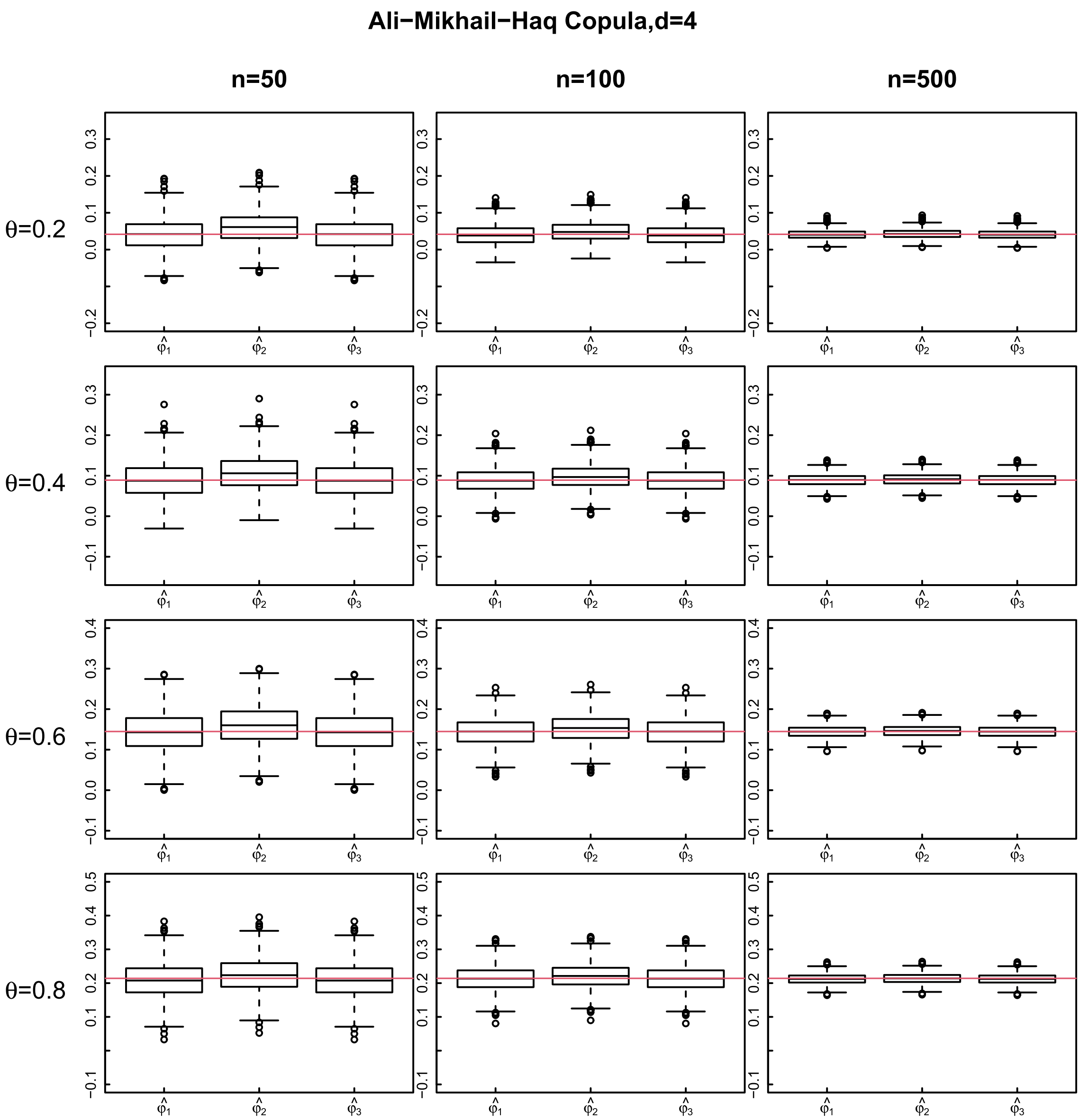

Fig. 3. Empirical distribution of three estimators of Spearman's footrule coefficient for the Ali-Mikhail-Haq d-copula with $d = 4$, parameter values $\theta = \{0.2, 0.4, 0.6, 0.8\}$ and sample sizes $n = \{50, 100, 500\}$. Horizontal red line in each graph sets the true value of Spearman's footrule.

where $\varphi_d'(0)$ and $\mu_V'(0)$ are the derivatives with respect to θ of the asymptotic means $\varphi_d(\theta)$ and $\mu_V(\theta)$, respectively, evaluated at $\theta = 0$, i.e., $\mu_V'(0) = (\partial\mu_V(\theta)/\partial\theta)|_{\theta=0}$ and $\varphi_d'(0) = (\partial\varphi_d(\theta)/\partial\theta)|_{\theta=0}$, and $\sigma_\varphi^2(0)$ and $\sigma_V^2(0)$ stand for the asymptotic variances of $\widehat{\varphi}_{nd}$ and V_n, respectively, at independence. Note that the expression of $\sigma_\varphi^2(0)$ is given in equation (20), where it was denoted as σ_Π^2. If $ARE(\widehat{\varphi}_{nd}, V_n) > 1$, Spearman's footrule will be a locally more powerful test statistic for independence than V_n, whereas $ARE(\widehat{\varphi}_{nd}, V_n) < 1$ will indicate that V_n is more powerful.

As far as we know, the only paper comparing the merits of the tests based on Spearman's footrule with other test statistics for independence is Genest et al. [14], who confine their analysis to the bidimensional case and conclude that no general recommendation can be made, as the results depend on the specific class of alternatives. In a multivariate setting, Stepanova [37] establishes conditions for Pitman optimality of multivariate tests for independence based on multivariate versions of Kendall's tau and Spearman's rho. In particular, the author proves that the average

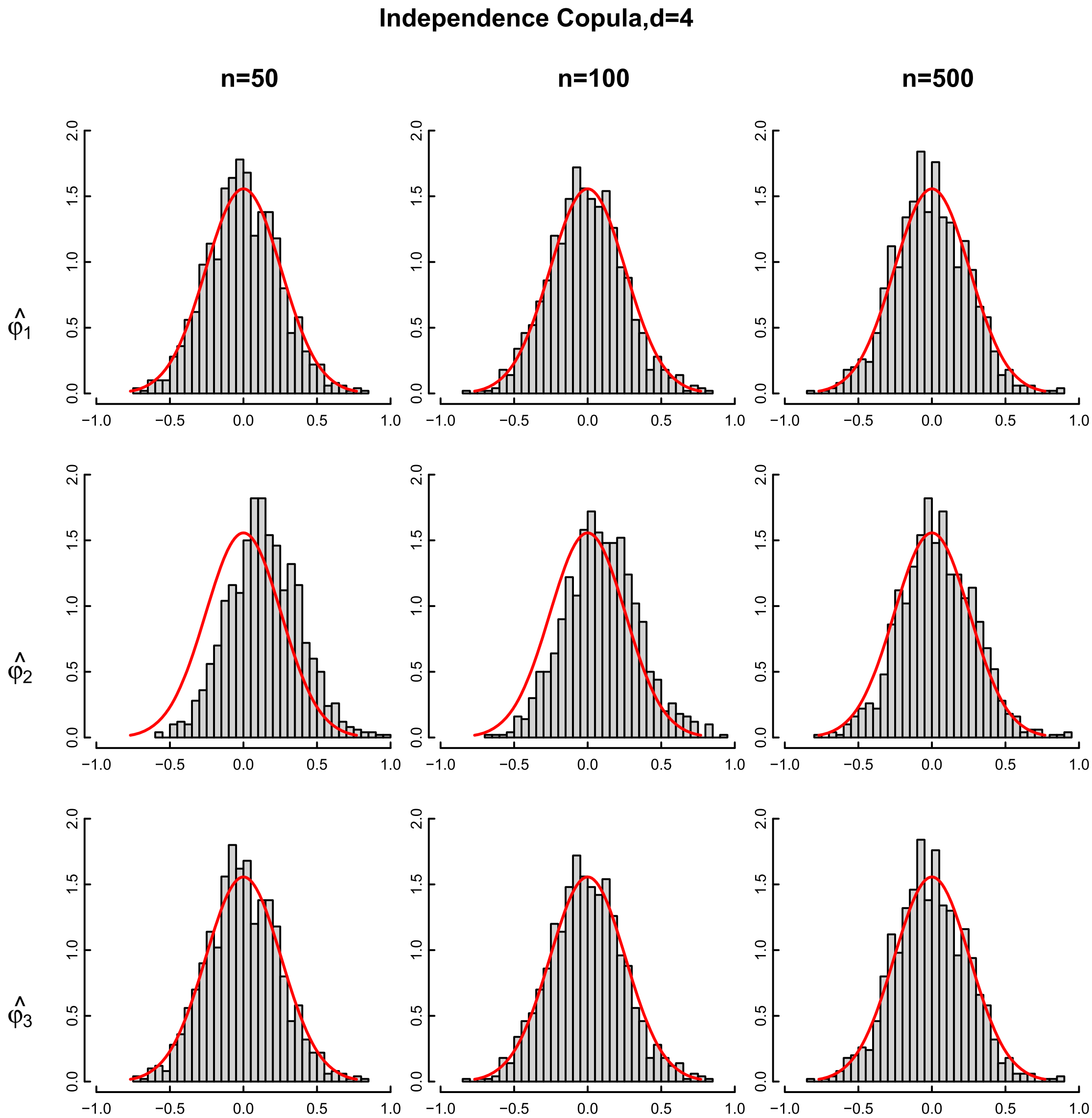

Fig. 4. Empirical distribution (rescaled by $\sqrt{n}$) of three estimators of Spearman's footrule coefficient for the independent d-copula with $d = 4$ and sample sizes $n = \{50, 100, 500\}$. The corresponding asymptotic Normal distribution under independence is also displayed.

pairwise Kendall's tau and the average pairwise Spearman's rho are asymptotically equivalent (in the Pitman's sense) and they are Pitman optimal for some particular copulas, whereas for other alternatives, the multivariate version of Spearman's rho proposed by Joe [19] is Pitman's optimal. Quessy [31] complements the findings in Stepanova [37] by comparing eight Spearman-type statistics for independence under different copula models and also concludes that their performance can vary depending on the kind of local alternatives encountered.

Following this literature, we investigate Pitman's ARE of multivariate Spearman's footrule with respect to multivariate versions of Kendall's tau and Spearman's rho. In particular, we consider the multivariate coefficients in Nelsen [24], which can be written as follows:

$$\tau_d = \tau_d(\theta) = \frac{1}{2^{d-1}-1}\left[2^d\left(1-\int_0^1 K(t,\theta)\,\mathrm{d}t\right)-1\right], \tag{31}$$

$$\rho_d^+ = \rho_d^+(\theta) = \frac{d+1}{2^d - d - 1}\left[2^d \int\limits_{[0,1]^d} \Pi(\mathbf{u})\,\mathrm{d}C_\theta(\mathbf{u}) - 1\right],$$

$$\rho_d^- = \rho_d^-(\theta) = \frac{d+1}{2^d - d - 1}\left[2^d \int\limits_{[0,1]^d} C_\theta(\mathbf{u})\,\mathrm{d}\Pi(\mathbf{u}) - 1\right], \tag{32}$$

where C_θ is a one-parameter d-dimensional copula, with $d \geq 2$, and $K(t,\theta)$ is the so-called Kendall's distribution function of C_θ; see Barbe et al. [2] and Genest et al. [16]. Nelsen [24] notes that, when $d = 3$, τ_3 coincides with the average of the three possible pairwise Kendall's tau. For further comparisons of our results with previous literature, we also consider another multivariate version of Spearman's rho, defined as the average of all possible pairwise Spearman's correlation coefficients, i.e.[2]

$$\rho_d^* = \rho_d^*(\theta) = \frac{2}{d(d-1)}\sum_{j<k}\left[12 \int\limits_{[0,1]^2} C_\theta(u_j, u_k)\,\mathrm{d}u_j \mathrm{d}u_k - 3\right].$$

The corresponding statistics for testing independence, based on a random sample $\mathbf{X}_i = (X_{i1}, ..., X_{id})$, $i = 1, \ldots, n$, of a continuous random vector $\mathbf{X}$ with copula C_θ, are the following:

$$\widehat{\tau}_{nd} = \frac{1}{2^{d-1} - 1}\left[-1 + \frac{2^d}{n(n-1)}\sum_{i \neq j} \mathbf{1}\{\mathbf{X}_i \leq \mathbf{X}_j\}\right],$$

$$\widehat{\rho}_{nd}^{+} = \left[\frac{1}{n}\sum_{i=1}^{n}\prod_{j=1}^{d} R_{ij} - \left(\frac{n+1}{2}\right)^d\right] \Big/ \left[\frac{1}{n}\sum_{i=1}^{n} i^d - \left(\frac{n+1}{2}\right)^d\right],$$

$$\widehat{\rho}_{nd}^{-} = \left[\frac{1}{n}\sum_{i=1}^{n}\prod_{j=1}^{d} \overline{R}_{ij} - \left(\frac{n+1}{2}\right)^d\right] \Big/ \left[\frac{1}{n}\sum_{i=1}^{n} i^d - \left(\frac{n+1}{2}\right)^d\right],$$

$$\widehat{\rho}_{nd}^{*} = \frac{12}{n^2 - 1}\left[\frac{2}{d(d-1)}\sum_{1 \leq j < k \leq d} \frac{1}{n}\sum_{i=1}^{n} R_{ij}R_{ik} - \left(\frac{n+1}{2}\right)^2\right],$$

where R_{ij} are the ranks defined in Section 3 and $\overline{R}_{ij} = n + 1 - R_{ij}$. The statistic $\widehat{\tau}_{nd}$ is taken from Genest et al. [15], whereas the statistic $\widehat{\rho}_{nd}^{+}$ is in Joe [19], Stepanova [37] and Pérez and Prieto-Alaiz [29], the statistic $\widehat{\rho}_{nd}^{-}$ is taken from Pérez and Prieto-Alaiz [29] and $\widehat{\rho}_{nd}^{*}$ is taken from Joe [19] and Stepanova [37]. These last three statistics are slightly different to those in Quessy [31] and Schmid and Schmidt [32], but they are asymptotically equivalent. Noticeably, for $d = 3$, the statistic $\widehat{\tau}_{n3}$ becomes the average of the three possible empirical pairwise Kendall's tau.

Under certain regularity conditions, Genest et al. [15] show that the statistic $\sqrt{n}\,(\widehat{\tau}_{nd} - \tau_d)$ is an asymptotically zero-mean normal variable whose limiting variance, under the null hypothesis of independence, i.e., when $C_0 = \Pi$, is

$$\sigma_\tau^2(0) = \frac{2}{(2^{d-1} - 1)^2}\left[\left(\frac{4}{3}\right)^d + \left(\frac{2}{3}\right)^d - 2\right]. \tag{33}$$

Similarly, it can be proved (see Quessy [31], Schmid and Schmidt [32] and Stepanova [37]) that the statistics $\sqrt{n}\,(\widehat{\rho}_{nd}^{+} - \rho_d^+)$, $\sqrt{n}\,(\widehat{\rho}_{nd}^{-} - \rho_d^-)$ and $\sqrt{n}\,(\widehat{\rho}_{nd}^{*} - \rho_d^*)$ are asymptotically zero-mean normal variables, whose limiting variances under independence are:

$$\sigma_{\rho^-}^2(0) = \sigma_{\rho^+}^2(0) = \frac{(d+1)^2}{(2^d - d - 1)^2}\left[\left(\frac{4}{3}\right)^d - \frac{d}{3} - 1\right], \tag{34}$$

$$\sigma_{\rho^*}^2(0) = \frac{2}{d(d-1)}. \tag{35}$$

2 Notice that Quessy [31] and Schmid and Schmidt [32] denote the coefficients ρ_d^-, ρ_d^+ and ρ_d^* as $\rho_{1,d}$, $\rho_{2,d}$ and $\rho_{3,d}$, respectively, whereas Joe [19] denoted ρ_d^- and ρ_d^+ as $\overline{\omega}$ and ω, respectively, and Wolff [40] introduced ρ_d^- as ρ_d.

Table 3
Asymptotic relative efficiency of multivariate Spearman's footrule with respect to multivariate Kendall's tau and Spearman's rho.

Model	FGM			AMH				
Statistic	$d=2$	$d=4$	$d=6$	$d=2$	$d=3$	$d=4$	$d=5$	$d=6$
$\widehat{\tau}_{nd}$	0.9	1.898	2.030	0.9	0.9	0.859	0.787	0.701
$\widehat{\rho}^{+}_{nd}$	0.9	2.313	2.869	0.9	0.810	0.705	0.597	0.496
$\widehat{\rho}^{-}_{nd}$	0.9	2.313	2.869	0.9	1.266	1.712	2.242	2.876
$\widehat{\rho}^{*}_{nd}$	0.9	–	–	0.9	0.9	0.875	0.831	0.780

Now, using equation (30), the Pitman's ARE of the empirical Spearman's footrule with respect to the empirical Kendall's tau, for instance, is computed as

$$ARE(\widehat{\varphi}_{nd}, \widehat{\tau}_{nd}) = \left(\frac{\varphi'_d(0)}{\sigma_\varphi(0)}\right)^2 \Big/ \left(\frac{\tau'_d(0)}{\sigma_\tau(0)}\right)^2,$$

where $\sigma^2_\tau(0)$ is given in (33) and $\tau'_d(0) = (\partial\tau_d(\theta)/\partial\theta)|_{\theta=0}$ will be calculated for each particular alternative. The Pitman's ARE of $\widehat{\varphi}_{nd}$ compared with $\widehat{\rho}^{-}_{nd}$, $\widehat{\rho}^{+}_{nd}$ and $\widehat{\rho}^{*}_{nd}$ are computed in a similar way.

To illustrate our results, we will chose as alternatives the multivariate FMG copula in Example 1 (notice that this generalization is different from that used in Quessy [31] and Stepanova [37]) and the multivariate AMH copula in Example 3. These two families are one-parameter d-copulas, say C_θ, that reduce to the independence copula when $\theta = 0$, i.e., $C_0 = \Pi$, so the hypothesis of independence can be regarded as $H_0 : \theta = 0$ in both cases. Moreover, the election of these families enables comparing our results with the existing literature mentioned above. Given below are the results for the cases $d = \{3, 4, 5, 6\}$. As a benchmark, we also compute the results for $d = 2$ (recall that, in this case, $\rho^-_2 = \rho^+_2 = \rho^*_2$ and $\widehat{\rho}^{-}_{n2} = \widehat{\rho}^{+}_{n2} = \widehat{\rho}^{*}_{n2}$).

Example 4. Let $C^{\mathbf{FGM}}_\theta$ be the FGM d-copula in equation (8). As shown in Úbeda-Flores [38], when d is odd, one gets $\tau_d = \varphi_d = 0$. Hence, for this family we will only perform comparisons for $d = \{2, 4, 6\}$. Moreover, for $d \geq 3$, all the margins of any dimension $j \geq 2$ are Π in the corresponding dimension j and so, the average pairwise Spearman's rho ρ^*_d equals 0. Hence, for $d \geq 3$, the corresponding statistic $\widehat{\rho}^{*}_{nd}$ will not be considered. Noticeably, for this family it is possible to work out closed-form expressions of the four coefficients to be analysed. In particular, the expression of multivariate Spearman's footrule is given in equation (9) and, from this equation, one easily deduces[3]

$$\varphi'_d\left(C^{\mathbf{FGM}}_0\right) = \frac{(1+(-1)^d)(d+1)(d!)^2}{(d-1)(2d+1)!}.$$

An explicit expression of multivariate Kendall's tau for this family is given in Genest et al. [15], from which it is obtained immediately

$$\tau'_d\left(C^{\mathbf{FGM}}_0\right) = \frac{(1+(-1)^d)}{3^d(2^{d-1}-1)}.$$

Finally, in Appendix E, we derive closed-form expressions of multivariate Spearman's ρ^+_d and ρ^-_d for this copula model. From these formulae, one deduces

$$\rho^{+\prime}_d\left(C^{\mathbf{FGM}}_0\right) = \frac{(-1)^d(d+1)}{3^d(2^d-d-1)}, \quad \rho^{-\prime}_d\left(C^{\mathbf{FGM}}_0\right) = \frac{(d+1)}{3^d(2^d-d-1)}.$$

Now, simple calculations based on the expressions above and those in equations (33)-(34), yield the results displayed in the first columns of Table 3. As expected, for $d = 2$, Spearman's footrule is less efficient than Kendall's tau and Spearman's rho, which are equivalent. Recall that the latter are optimal for the bivariate FGM; see Stepanova [37].

[3] To avoid possible confusion, $\varphi'_d\left(C^{\mathbf{FGM}}_0\right)$ will stand for the value of $\varphi'_d(0)$ in the FGM copula, $\tau'_d\left(C^{\mathbf{FGM}}_0\right)$ for the value of $\tau'_d(0)$ in the FGM copula, and so on. An equivalent notation will be used for the AMH copula.

However, for $d \geq 3$, the Spearman's footrule is always the best, while $\widehat{\tau}_{nd}$ is the second one and $\widehat{\rho}^{-}_{nd}$ and $\widehat{\rho}^{+}_{nd}$, which are asymptotically equivalent, are not worth considering for these local alternatives. Hence, for this family and $d \geq 3$, the statistics can be ordered in terms of their asymptotic efficiency, from the best to the worst, as

$$\widehat{\varphi}_{nd} > \widehat{\tau}_{nd} > \widehat{\rho}^{-}_{nd} = \widehat{\rho}^{+}_{nd}$$

Example 5. Let $C^{\mathbf{AMH}}_{\theta}$ be the AMH d-copula in equation (12). Unlike Example 4, in this case, there are no simple closed-form expressions of the coefficients to be analysed. For the Spearman's footrule, we do have an explicit formula in Equation (14), from which it is a routine exercise to check that

$$\varphi'_d\left(C^{\mathbf{AMH}}_0\right) = \frac{d}{(d+2)(2d+1)} + \frac{d+1}{d-1}\left[\sum_{i=2}^{d}(-1)^i\binom{d}{i}\frac{i(i-1)}{(i+1)(i+2)(2i+1)}\right].$$

For the other four coefficients, we directly focus on computing and evaluating their derivatives under $\theta = 0$. Regarding Kendall's tau, in Appendix E we derive a recursive general formula to compute $\tau'_d(0)$ for Archimedean d-copulas, using the results in Barbe et al. [2] and Genest et al. [16] based on the Kendall's distribution function. From this formula, the following particular values for the AMH come up

$$\tau'_2\left(C^{\mathbf{AMH}}_0\right) = \tau'_3\left(C^{\mathbf{AMH}}_0\right) = 0.222,\ \tau'_4\left(C^{\mathbf{AMH}}_0\right) = 0.194,\ \tau'_5\left(C^{\mathbf{AMH}}_0\right) = 0.156,\ \tau'_6\left(C^{\mathbf{AMH}}_0\right) = 0.120.$$

In turn, the derivatives of the multivariate Spearman's ρ^{+}_{d}, ρ^{-}_{d} and $\rho^{*\prime}_{d}$ for the AMH copula, evaluated at $\theta = 0$, are (see Quessy [31])

$$\rho^{+\prime}_d\left(C^{\mathbf{AMH}}_0\right) = \frac{(d+1)}{(2^d-d-1)}\left[\left(\frac{4}{3}\right)^d - \frac{d}{3} - 1\right],$$

$$\rho^{-\prime}_d\left(C^{\mathbf{AMH}}_0\right) = \frac{(d+1)}{(2^d-d-1)}\left[\left(\frac{2}{3}\right)^d + \frac{d}{3} - 1\right],$$

$$\rho^{*\prime}_d\left(C^{\mathbf{AMH}}_0\right) = \frac{1}{3}.$$

From these expressions above and those in (33)-(35), the results displayed in the last columns of Table 3 come up. One sees that, for $d = 2$, the results for the bivariate FGM and AMH copulas coincide. This does not come as a surprise since, in the bivariate case, $\varphi'_2\left(C^{\mathbf{AMH}}_0\right) = \varphi'_2\left(C^{\mathbf{FGM}}_0\right)$, $\tau'_2\left(C^{\mathbf{AMH}}_0\right) = \tau'_2\left(C^{\mathbf{FGM}}_0\right)$, and $\rho'_2\left(C^{\mathbf{AMH}}_0\right) = \rho'_2\left(C^{\mathbf{FGM}}_0\right)$. For $d \geq 3$, Spearman's footrule is always more efficient than $\widehat{\rho}^{-}_{nd}$, which becomes the poorest statistic for this kind of alternatives. By contrast, Spearman's footrule is always less efficient than Kendall's tau and Spearman's $\widehat{\rho}^{*}_{nd}$ and $\widehat{\rho}^{+}_{nd}$, and the latter always dominates the others. Keeping in mind that $\widehat{\rho}^{+}_{nd}$ is as its best under AMH kind of alternatives (see Quessy [31]), Spearman's footrule remains a good competitor, especially when d is small. As expected, for $d = 3$, Kendall's tau and Spearman's $\widehat{\rho}^{*}_{n3}$ become equivalent; recall that Stepanova [37] shows that average pairwise Spearman's and average pairwise Kendall's are asymptotically equivalent (in the Pitman sense). Hence, for this family and $d \geq 3$, the statistics can be ordered in terms of their asymptotic efficiency, as

$$\widehat{\rho}^{+}_{nd} > \widehat{\tau}_{nd} \geq \widehat{\rho}^{*}_{nd} > \widehat{\varphi}_{nd} > \widehat{\rho}^{-}_{nd}.$$

Other comparisons are not computed, but one can recover, for instance, the $ARE\left(\widehat{\rho}^{-}_{nd}, \widehat{\rho}^{+}_{nd}\right)$, via the relationship

$$ARE\left(\widehat{\rho}^{-}_{nd}, \widehat{\rho}^{+}_{nd}\right) = \frac{ARE\left(\widehat{\varphi}_{nd}, \widehat{\rho}^{+}_{nd}\right)}{ARE\left(\widehat{\varphi}_{nd}, \widehat{\rho}^{-}_{nd}\right)}.$$

In doing so, the values obtained for the AMH copula coincide with those in Quessy [31, Table 2], as expected.

6. Conclusions

In this paper, we propose two new nonparametric estimators of multivariate Spearman's footrule which are based on two sample versions of AOD measures of multivariate concordance and compare them with a previous estimator proposed in Úbeda-Flores [38]. We show that, one of the estimators proposed share some interesting features with the estimator in Úbeda-Flores [38]. In particular, they coincide in the bivariate case, where both reduce to the well-known empirical bivariate Spearman's footrule, whereas, in the trivariate case, they both become the average of the corresponding pairwise estimators. Moreover, we show that the three estimators analysed are asymptotically equivalent and we derive their asymptotic distribution under milder conditions and using a different approach than those in Genest et al. [14]. Furthermore, we show that, under independence, one of the new estimators proposed outperforms the others, since it is unbiased regardless of the dimensions considered. These results are further illustrated with Monte Carlo experiments based on the independent copula and two well-known one parameter copulas, namely a multivariate FGM copula and a multivariate Archimedean copula. These experiments also reveal that, in small samples, the performance of the three estimators considered is different. In particular, one of the two estimators proposed is nearly equivalent to the estimator in Úbeda-Flores [38] and clearly outperforms the others, especially in small samples. The paper also investigates the performance of this estimator, in terms of their Pitman asymptotic relative efficiency, as compared to Spearman's rho and Kendall's tau, when they are used as statistics for testing multivariate independence. As one could expect, the results vary considerably depending on the kind of alternatives encountered: Spearman's footrule seems to be locally optimal for the multivariate FGM alternative when $d \geq 3$, but it is less efficient when the alternative is the multivariate AMH, where Spearman's $\widehat{\rho}_{nd}^{+}$ and $\widehat{\rho}_{nd}^{-}$ seem to be the most and least efficient, respectively. Further comparisons with other multivariate measures of association, like Gini's gamma or Blomqvist beta, would be interesting for future works.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgements

The third author is partially supported by the PID2020-113350GB-I00 grant of the Ministerio de Ciencia e Innovación (Spain) and by the "Severo Ochoa Programme for Centres of Excellence in R&D" (CEX2019-000904-S). The fifth author acknowledges the support of the program FEDER-Andalucía 2014-2020 under research project UAL2020-AGR-B1783 and project PID2021-122657OB-I00 by the Ministerio de Ciencia e Innovación (Spain).

Appendix A

Lemma 1. *Let* $f(t) = \frac{1-\theta}{e^t-\theta}$ *for all* $t \geq 0$, *with* $\theta \in [0, 1[$. *Then*

$$(-1)^k \frac{d^k}{dt^k} f(t) = \frac{(1-\theta)e^t \sum_{i=0}^{k-1} p_i \theta^i e^{(k-i-1)t}}{(e^t-\theta)^{k+1}} \tag{36}$$

for all $t \in]0, +\infty[$, *where* $p_i > 0$ *for all* $i = 0, 1, \ldots, k-1$, *with* $p_0 = p_{k-1} = 1$.

Proof. We prove this result by induction on k. If $k = 1$, then

$$(-1)\frac{d}{dt} f(t) = \frac{(1-\theta)e^t p_0}{(e^t-\theta)^2}$$

for all $t > 0$. Now, assume the result is true for any value $2, \ldots, k$. Then, by using (36), we have

$$
\begin{aligned}
&-\frac{d}{dt}\left((-1)^k \frac{d^k}{dt^k}\right) f(t) \\
&= -(1-\theta)e^t \frac{\left(\sum_{i=0}^{k-1} p_i\theta^i e^{(k-i-1)t} + \sum_{i=0}^{k-2} p_i\theta^i (k-i-1)e^{(k-i-1)t}\right)(e^t-\theta) - \sum_{i=0}^{k-1} p_i\theta^i e^{(k-i-1)t}(k+1)e^t}{(e^t-\theta)^{k+2}} \\
&= -(1-\theta)e^t \frac{(e^t-\theta)\sum_{i=0}^{k-1} p_i\theta^i (k-i)e^{(k-i-1)t} - (k+1)\sum_{i=0}^{k-1} p_i\theta^i e^{(k-i)t}}{(e^t-\theta)^{k+2}} \\
&= -(1-\theta)e^t \frac{\sum_{i=0}^{k-1} p_i\theta^i (k-i)e^{(k-i)t} - \sum_{i=1}^{k} p_{i-1}\theta^i (k-i+1)e^{(k-i)t} - \sum_{i=0}^{k-1} p_i\theta^i (k+1)e^{(k-i)t}}{(e^t-\theta)^{k+2}} \\
&= \frac{(1-\theta)e^t \sum_{i=0}^{k} q_i\theta^i e^{(k-i)t}}{(e^t-\theta)^{k+2}}
\end{aligned}
$$

for all $t > 0$, with $q_0 = p_0 = 1$, $q_k = p_{k-1} = 1$ and $q_i = (i+1)p_i + (k-i+1)p_{i-1} > 0$ for every $i = 1, \ldots, k-1$, whence the result follows. □

Appendix B. Proof of Theorem 1

To prove Theorem 1 we need additional notation and some preliminary results. We denote the empirical distribution function of the d-dimensional vector $\mathbf{X} = (X_1, \ldots, X_d)$ by H_n, defined as

$$H_n(\mathbf{x}) = \frac{1}{n+1}\sum_{i=1}^{n} \mathbf{1}\{\mathbf{X}_i \leq \mathbf{x}\}$$

for $\mathbf{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d$, where $\mathbf{X}_i$, for $i = 1, \ldots, n$, is a sample of n independent random vectors from $\mathbf{X}$, and $\mathbf{X}_i \leq \mathbf{x}$ means $X_{ij} \leq x_j$, for all $j = 1, \ldots, d$, and we denote the empirical marginal distribution of X_{ij} by F_{nj}, defined as

$$F_{nj}(z) = \frac{1}{n+1}\sum_{i=1}^{n} \mathbf{1}\left\{X_{ij} \leq z\right\}.$$

Note that in the case of data without ties, we have

$$F_{nj}\left(X_{ij}\right) = \frac{R_{ij}}{n+1},$$

where R_{ij} denotes the *rank* of X_{ij} among $\{X_{1j}, \ldots, X_{nj}\}$, with $j = 1, \ldots, d$.

Consider

$$A_n = \frac{1}{n}\sum_{i=1}^{n} M(F_{n1}(X_{i1}), \ldots, F_{nd}(X_{id})) - \int_{[0,1]^d} M(\mathbf{u})\, \mathrm{d}C(\mathbf{u}), \tag{37}$$

$$B_n = \frac{1}{n}\sum_{i=1}^{n} \overline{M}(F_{n1}(X_{i1}), \ldots, F_{nd}(X_{id})) - \int_{[0,1]^d} \overline{M}(\mathbf{u})\, \mathrm{d}C(\mathbf{u}). \tag{38}$$

Now we have

$$\widehat{Q}'_{nd}(M, C) - Q'_d(M, C) = A_n + B_n, \tag{39}$$

and

$$A_n = A_n^* + \frac{1}{n}\sum_{i=1}^{n}\left(M(F_{n1}(X_{i1}),\dots,F_{nd}(X_{id})) - M(F_1(X_{i1}),\dots,F_d(X_{id}))\right),$$

$$B_n = B_n^* + \frac{1}{n}\sum_{i=1}^{n}\left(\overline{M}(F_{n1}(X_{i1}),\dots,F_{nd}(X_{id})) - \overline{M}(F_1(X_{i1}),\dots,F_d(X_{id}))\right),$$

where

$$A_n^* = \frac{1}{n}\sum_{i=1}^{n}\left(M(F_1(X_{i1}),\dots,F_d(X_{id})) - \mathbb{E}M(F_1(X_{i1}),\dots,F_d(X_{id}))\right), \tag{40}$$

$$B_n^* = \frac{1}{n}\sum_{i=1}^{n}\left(\overline{M}(F_1(X_{i1}),\dots,F_d(X_{id})) - \mathbb{E}\overline{M}(F_1(X_{i1}),\dots,F_d(X_{id}))\right). \tag{41}$$

We next give an auxiliary proposition, where ω denotes the elementary event and $\omega \to n_o(\omega)$ is a random number.

Proposition 1. *Assume that the (univariate) distributions $F_1,\dots,F_d$ are continuous. Then*

a) $\max_{j=1\dots d}\sup_{x\in\mathbb{R}}\left|F_{jn}(x) - F_j(x)\right| = O_{\mathbb{P}}(n^{-1/2})$, *and*

b) $\max_{j=1\dots d}\sup_{x\in\mathbb{R}}\left|F_{jn}(x) - F_j(x)\right| \le \kappa_1\sqrt{\frac{\ln\ln n}{n}}$ *a.s.*

for $n \ge n_0(\omega)$ with a constant $\kappa_1 > \frac{1}{2}\sqrt{2}$.

Proof. Assertion a) is a consequence of the Dvoretzky-Kiefer-Wolfowitz inequality (see van der Vaart [39], p.268):

$$\mathbb{P}\left\{\sqrt{n}\sup_{x\in\mathbb{R}}\left|F_{jn}(x) - F_j(x)\right| > t\right\} \le 2e^{-2t^2}$$

for $n, t > 0$, $j = 1,\dots,d$.

Assertion b) follows from van der Vaart [39], p. 268, for example. □

Lemma 2. *Consider the expressions (37), (38), (40) and (41), and suppose that Assumption $\mathcal{A}_C$ is fulfilled. Then we have*

i. $A_n - A_n^* = D_n + o_{\mathbb{P}}\left(n^{-1/2}\right)$, *where*

$$D_n = \sum_{j=1}^{d}\frac{1}{n}\sum_{i=1}^{n}\left(F_{nj}(X_{ij}) - F_j(X_{ij})\right)\mathbf{1}\left\{F_j(X_{ij}) < F_l(X_{il})\ \forall l \ne j\right\}.$$

ii. $B_n - B_n^* = E_n + o_{\mathbb{P}}\left(n^{-1/2}\right)$, *where*

$$E_n = \sum_{j=1}^{d}\frac{1}{n}\sum_{i=1}^{n}\left(-F_{nj}(X_{ij}) + F_j(X_{ij})\right)\mathbf{1}\left\{F_j(X_{ij}) > F_l(X_{il})\ \forall l \ne j\right\}.$$

Proof. Firstly, note that $\mathbf{1}(A\cap B) = \mathbf{1}(B) - \mathbf{1}(\overline{A}\cap B)$ for any two events A, B, and

$$\mathbb{P}\left\{F_j(X_{ij}) = F_l(X_{il})\right\} = 0 \ \text{ for } l \ne j.$$

We prove part i. We obtain

$$\begin{aligned} A_n - A_n^* = & \sum_{j,k=1}^{d} \frac{1}{n} \sum_{i=1}^{n} \left(F_{nk}(X_{ik}) - F_j(X_{ij})\right) \\ & \quad \mathbf{1}\Big\{F_{nk}(X_{ik}) < F_{nl}(X_{il}) \ \forall l < k, F_{nk}(X_{ik}) \le F_{nl}(X_{il}) \ \forall l > k, \\ & \quad\quad F_j(X_{ij}) < F_m(X_{im}) \ \forall m \ne j\Big\} \\ = & \ D_n + D_{n1} - D_{n2} \ \ a.s., \end{aligned} \tag{42}$$

where

$$\begin{aligned} D_{n1} = & \sum_{j=1}^{d} \sum_{\substack{1 \le k \le d \\ k \ne j}} \frac{1}{n} \sum_{i=1}^{n} \left(F_{nk}(X_{ik}) - F_j(X_{ij})\right) \\ & \mathbf{1}\Big\{F_{nk}(X_{ik}) < F_{nl}(X_{il}) \ \forall l < k, F_{nk}(X_{ik}) \le F_{nl}(X_{il}) \ \forall l > k, \\ & \quad F_j(X_{ij}) < F_l(X_{il}) \ \forall l \ne j\Big\}, \\ D_{n2} = & \sum_{j=1}^{d} \frac{1}{n} \sum_{i=1}^{n} \left(F_{nj}(X_{ij}) - F_j(X_{ij})\right) \\ & \mathbf{1}\left\{\max\{(F_{nj}(X_{ij}) \ge F_{nl}(X_{il}) \text{ for some } l < j), (F_{nj}(X_{ij}) > F_{nl}(X_{il}) \text{ for some } l > j)\},\right. \\ & \quad \left. F_j(X_{ij}) < F_m(X_{im}) \ \forall m \ne j\right\}. \end{aligned}$$

Define

$$\gamma_n := \max_{j=1\dots d} \sup_{x \in \mathbb{R}} \left|F_{jn}(x) - F_j(x)\right| \ \text{ and } \ \bar{\gamma}_n := \kappa_1 \sqrt{\frac{\ln \ln n}{n}},$$

with κ_1 as in Proposition 1. For $F_{nk}(X_{ik}) \le F_{nj}(X_{ij}),\ F_j(X_{ij}) < F_k(X_{ik})$, we have

$$\begin{aligned} & F_{nk}(X_{ik}) - F_j(X_{ij}) \le F_{nj}(X_{ij}) - F_j(X_{ij}) \le \gamma_n, \\ & F_{nk}(X_{ik}) - F_j(X_{ij}) > F_{nk}(X_{ik}) - F_k(X_{ik}) \ge -\gamma_n \ \ a.s., \end{aligned} \tag{43}$$

whenever $k \ne j$. On the other hand,

$$F_k(X_{ik}) - \gamma_n \le F_{nk}(X_{ik}) \le F_{nj}(X_{ij}) \le F_j(X_{ij}) + \gamma_n < F_k(X_{ik}) + \gamma_n \quad a.s. \tag{44}$$

Hence

$$\begin{aligned} |D_{n1}| & \le \gamma_n \sum_{j=1}^{d} \sum_{\substack{1 \le k \le d \\ k \ne j}} \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}\left\{F_{nk}(X_{ik}) \le F_{nj}(X_{ij}),\ F_j(X_{ij}) < F_k(X_{ik})\right\} \\ & \le \gamma_n \sum_{j=1}^{d} \sum_{\substack{1 \le k \le d \\ k \ne j}} \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}\left\{F_k(X_{ik}) - 2\gamma_n \le F_j(X_{ij}) < F_k(X_{ik})\right\} \\ & \le \gamma_n \sum_{j=1}^{d} \sum_{\substack{1 \le k \le d \\ k \ne j}} \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}\left\{-2\bar{\gamma}_n \le F_j(X_{ij}) - F_k(X_{ik}) < 0\right\} \ \ a.s. \end{aligned}$$

Applying the law of iterated logarithm for empirical processes (cf. Van der Vaart [39], p. 268) on the sample $\{F_j(X_{ij}) - F_k(X_{ik})\}_{i=1\dots n}$, it follows that

$$|D_{n1}| \leq \gamma_n \left(O(\bar{\gamma}_n) + \sum_{j=1}^{d} \sum_{\substack{1\leq k\leq d \\ k\neq j}} \mathbb{P}\left\{-2\bar{\gamma}_n \leq F_j(X_{ij}) - F_k(X_{ik}) < 0\right\} \right) \quad a.s.$$

Let us fix $j, k, j \neq k$, and $U_i = F_j\left(X_{ij}\right)$ and $V_i = F_k\left(X_{ik}\right)$. Therefore, by assumption $\mathcal{A}_C$, it follows that $C_{jk}^{(1)}$ is uniformly continuous on a neighbourhood of the diagonal and

$$\begin{aligned}\mathbb{P}\left\{-2\bar{\gamma}_n \leq F_j(X_{ij}) - F_k(X_{ik}) < 0\right\} &= \int_0^1 \mathbb{P}\left\{-2\bar{\gamma}_n \leq U_i - V_i < 0 \mid U_i = u\right\} \mathrm{d}u \\ &= \int_0^1 \left(C_{jk}^{(1)}(u, u+2\bar{\gamma}_n) - C_{jk}^{(1)}(u,u)\right) \mathrm{d}u \\ &= o(1),\end{aligned}$$

and therefore,

$$|D_{n1}| = o_{\mathbb{P}}\left(n^{-1/2}\right). \tag{45}$$

Now notice that $\left(\bigcup_{l\neq j} E_l\right) \cap G \subset \bigcup_{l\neq j}(E_l \cap G_l)$ for events $E_1, \ldots, E_d, G_1, \ldots, G_d$, and where $G = \bigcap_{m=1}^{d} G_m$. Then an application of Proposition 1 leads to

$$\begin{aligned}|D_{n2}| &\leq \gamma_n \sum_{j=1}^{d} \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}\left\{F_{nj}(X_{ij}) \geq F_{nl}(X_{il}) \text{ for some } l \neq j,\ F_j(X_{ij}) < F_m(X_{im})\ \forall m \neq j\right\} \\ &\leq \gamma_n \sum_{j=1}^{d} \sum_{\substack{1\leq l\leq d \\ l\neq j}} \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}\left\{F_{nj}(X_{ij}) \geq F_{nl}(X_{il}),\ F_j(X_{ij}) < F_l(X_{il})\right\} \\ &= o_{\mathbb{P}}\left(n^{-1/2}\right),\end{aligned} \tag{46}$$

the latter identity analogously to the considerations on D_{n1}. Thus, part i. of the lemma is a consequence of (42)–(46).

Now we prove part ii. We obtain

$$\begin{aligned}B_n - B_n^* = &\sum_{j,k=1}^{d} \frac{1}{n} \sum_{i=1}^{n} \left(-F_{nk}(X_{ik}) + F_j(X_{ij})\right) \\ &\mathbf{1}\left\{F_{nk}(X_{ik}) > F_{nl}(X_{il})\ \forall l < k, F_{nk}(X_{ik}) \geq F_{nl}(X_{il})\ \forall l > k,\ F_j(X_{ij}) > F_m(X_{im})\ \forall m \neq j\right\} \\ = &E_n + E_{n1} - E_{n2} \quad a.s.,\end{aligned} \tag{47}$$

where

$$\begin{aligned}E_{n1} = &\sum_{j=1}^{d} \sum_{\substack{1\leq k\leq d \\ k\neq j}} \frac{1}{n} \sum_{i=1}^{n} \left(-F_{nk}(X_{ik}) + F_j(X_{ij})\right) \\ &\mathbf{1}\left\{F_{nk}(X_{ik}) > F_{nl}(X_{il})\ \forall l < k, F_{nk}(X_{ik}) \geq F_{nl}(X_{il})\ \forall l > k, F_j(X_{ij}) > F_m(X_{im})\ \forall m \neq j\right\}, \\ E_{n2} = &\sum_{j=1}^{d} \frac{1}{n} \sum_{i=1}^{n} \left(-F_{nj}(X_{ij}) + F_j(X_{ij})\right) \\ &\mathbf{1}\left\{\max\{(F_{nj}(X_{ij}) \leq F_{nl}(X_{il}) \text{ for some } l < j)(F_{nj}(X_{ij}) < F_{nl}(X_{il}) \text{ for some } l > j)\},\right. \\ &\qquad \left. F_j(X_{ij}) > F_m(X_{im})\ \forall m \neq j\right\}.\end{aligned}$$

For $F_{nk}(X_{ik}) \geq F_{nj}(X_{ij})$ and $F_j(X_{ij}) > F_k(X_{ik})$, it follows that

$$-F_{nk}(X_{ik}) + F_j(X_{ij}) \leq -F_{nj}(X_{ij}) + F_j(X_{ij}) \leq \gamma_n,$$
$$-F_{nk}(X_{ik}) + F_j(X_{ij}) > -F_{nk}(X_{ik}) + F_k(X_{ik}) \geq -\gamma_n,$$

whenever $k \neq j$. On the other hand,

$$F_k(X_{ik}) + \gamma_n \geq F_{nk}(X_{ik}) \geq F_{nj}(X_{ij}) \geq F_j(X_{ij}) - \gamma_n > F_k(X_{ik}) - \gamma_n.$$

Taking (43) and (44) into account and applying Proposition 1, we derive

$$\begin{aligned}|E_{n1}| &\leq \gamma_n \sum_{j=1}^{d} \sum_{\substack{1\leq k\leq d\\ k\neq j}} \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}\left\{F_{nk}(X_{ik}) \geq F_{nj}(X_{ij}),\ F_j(X_{ij}) > F_k(X_{ik})\right\} \\ &\leq \gamma_n \sum_{j=1}^{d} \sum_{\substack{1\leq k\leq d\\ k\neq j}} \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}\left\{F_j(X_{ij}) - 2\gamma_n \leq F_k(X_{ik}) < F_j(X_{ij})\right\}.\end{aligned}$$

Analogously to the considerations in part i., we obtain

$$|E_{n1}| = o_{\mathbb{P}}\left(n^{-1/2}\right). \tag{48}$$

By Proposition 1 and considerations in part i., we have

$$\begin{aligned}|E_{n2}| &\leq \gamma_n \sum_{j=1}^{d} \sum_{\substack{1\leq l\leq d\\ l\neq j}} \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}\left\{F_{nj}(X_{ij}) \leq F_{nl}(X_{il}),\ F_j(X_{ij}) > F_l(X_{il})\right\} \\ &= o_{\mathbb{P}}\left(n^{-1/2}\right). \end{aligned} \tag{49}$$

Identities (47)–(49) lead to part ii. of the lemma, and this completes the proof. □

Let V_n be a U-statistic with symmetric real-valued kernel function Λ (symmetric means that interchanging the arguments does not affect the values of Λ):

$$V_n = \binom{n}{2}^{-1} \sum_{1\leq i_1<i_2\leq n} \Lambda\left(\mathbf{Y}_{i_1}, \mathbf{Y}_{i_2}\right),$$

where $\mathbf{Y}_1, \ldots, \mathbf{Y}_n$ is the sample of i.i.d. random vectors. Define

$$\begin{aligned}\lambda(\mathbf{y}) &:= \mathbb{E}\Lambda(\mathbf{y}, \mathbf{Y}_2), \\ \zeta_1 &:= \mathbb{V}(\lambda(\mathbf{Y}_1)), \\ \beta &:= \mathbb{E}\lambda(\mathbf{Y}_1) = \mathbb{E}\Lambda\left(\mathbf{Y}_1, \mathbf{Y}_2\right),\end{aligned}$$

where $\mathbb{V}$ is the symbol for the variance. Note that $\lambda(\mathbf{y})$ is the conditional expectation of Λ given one argument equals $\mathbf{y}$.

In the next proposition we recall the central limit theorem for non-degenerate U-statistics (Theorem 5.5.1A in Serfling [34]).

Proposition 2. *Assume that Λ is symmetric and $\zeta_1 \neq 0$. Then*

$$\sqrt{n}\left(V_n - \beta\right) \xrightarrow{d} \mathcal{N}\left(0, 4\,\zeta_1\right).$$

Now we are in conditions to prove Theorem 1.

Proof of Theorem 1: Since

$$D_n + E_n = \frac{1}{n(n+1)} \sum_{i,k=1}^{n} \sum_{j=1}^{d} \left(\mathbf{1}\left\{X_{kj} \le X_{ij}\right\} - F_j\left(X_{ij}\right)\right) \mathbf{1}\left\{F_j\left(X_{ij}\right) < F_l\left(X_{il}\right) \forall l \neq j\right\}$$

$$- \frac{1}{n(n+1)} \sum_{i,k=1}^{n} \sum_{j=1}^{d} \left(\mathbf{1}\left\{X_{kj} \le X_{ij}\right\} - F_j\left(X_{ij}\right)\right) \mathbf{1}\left\{F_j\left(X_{ij}\right) > F_l\left(X_{il}\right) \forall l \neq j\right\}$$

$$= \frac{1}{n^2} \sum_{i,k=1}^{n} \sum_{j=1}^{d} \left(\mathbf{1}\left\{X_{kj} \le X_{ij}\right\} - F_j\left(X_{ij}\right)\right)$$

$$\cdot \left(\mathbf{1}\left\{F_j\left(X_{ij}\right) < F_l\left(X_{il}\right) \forall l \neq j\right\} - \mathbf{1}\left\{F_j\left(X_{ij}\right) > F_l\left(X_{il}\right) \forall l \neq j\right\}\right) + \breve{R}_n,$$

where

$$\left|\breve{R}_n\right| \le n^2 d \left(\frac{1}{n^2} - \frac{1}{n(n+1)}\right) = \frac{d}{n+1},$$

as a consequence of (39) and Lemma 2, we obtain

$$\widehat{Q}'_{nd}(M, C) - Q'_d(M, C) = A_n^* + B_n^* + D_n + E_n + o_{\mathbb{P}}\left(n^{-1/2}\right)$$

$$= A_n^* + B_n^*$$

$$+ \frac{1}{n^2} \sum_{i,k=1}^{n} \sum_{j=1}^{d} \left(\mathbf{1}\left\{X_{kj} \le X_{ij}\right\} - F_j(X_{ij})\right) \mathbf{1}\left\{F_j(X_{ij}) < F_l(X_{il}) \forall l \neq j\right\}$$

$$- \frac{1}{n^2} \sum_{i,k=1}^{n} \sum_{j=1}^{d} \left(\mathbf{1}\left\{X_{kj} \le X_{ij}\right\} - F_j(X_{ij})\right) \mathbf{1}\left\{F_j(X_{ij}) > F_l(X_{il}) \forall l \neq j\right\}$$

$$+ o_{\mathbb{P}}\left(n^{-1/2}\right)$$

$$= \frac{1}{n^2} \sum_{i=1}^{n} \sum_{k=1}^{n} \Lambda\left(\mathbf{X}_i, \mathbf{X}_k\right) - Q'_d(M, C) + o_{\mathbb{P}}\left(n^{-1/2}\right),$$

where

$$\Lambda(\mathbf{x}, \mathbf{y}) = M(F_1(x_1), \ldots, F_d(x_d)) + \overline{M}(F_1(x_1), \ldots, F_d(x_d))$$
$$+ \sum_{j=1}^{d} \left(\mathbf{1}\left\{y_j \le x_j\right\} - F_j(x_j)\right) \left(\mathbf{1}\left\{F_j(x_j) < F_l(x_l) \forall l \neq j\right\} - \mathbf{1}\left\{F_j(x_j) > F_l(x_l) \forall l \neq j\right\}\right)$$

for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$. Since

$$\frac{1}{n^2} \sum_{i=1}^{n} \Lambda(\mathbf{X}_i, \mathbf{X}_i) = O_{\mathbb{P}}\left(n^{-1}\right),$$

we can conclude

$$\widehat{Q}'_{nd}(M, C) - Q'_d(M, C) = V_n - Q'_d(M, C) + o_{\mathbb{P}}\left(n^{-1/2}\right),$$

where

$$V_n = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{k=i+1}^{n} \left(\Lambda(\mathbf{X}_i, \mathbf{X}_k) + \Lambda(\mathbf{X}_k, \mathbf{X}_i)\right).$$

V_n is a U-statistic with symmetric kernel $(\mathbf{x}, \mathbf{y}) \longmapsto \Lambda(\mathbf{x}, \mathbf{y}) + \Lambda(\mathbf{y}, \mathbf{x})$. Note that

$$h(\mathbf{x}) := \mathbb{E}\left(\Lambda\left(\mathbf{X}_i, \mathbf{x}\right) + \Lambda\left(\mathbf{x}, \mathbf{X}_i\right)\right)$$
$$= \sum_{j=1}^{d} \int_{\mathbb{R}^d} \left(\mathbf{1}\left\{x_j \le z_j\right\} - F_j(z_j)\right)\left(\mathbf{1}\left\{F_j(z_j) < F_l(z_l)\ \forall l \ne j\right\} - \mathbf{1}\left\{F_j(z_j) > F_l(z_l)\ \forall l \ne j\right\}\right) \mathrm{d}H(\mathbf{z})$$
$$+ M(F_1(x_1), \ldots, F_d(x_d)) + \overline{M}(F_1(x_1), \ldots, F_d(x_d)) + Q'_d(M, C).$$

By applying Proposition 2, for $\lambda(.) = h(.)/2$, $\beta = Q'_d(M, C)$ and $\zeta_1 = \sigma^2$, and taking into account

$$\sigma^2 = \mathbb{V}(h(\mathbf{X}_i))$$
$$= \int_{\mathbb{R}^d} \Bigg(\sum_{j=1}^{d} \int_{\mathbb{R}^d} \left(\mathbf{1}\left\{x_j \le z_j\right\} - F_j(z_j)\right)$$
$$\left(\mathbf{1}\left\{F_j(z_j) < F_l(z_l)\ \forall l \ne j\right\} - \mathbf{1}\left\{F_j(z_j) > F_l(z_l)\ \forall l \ne j\right\}\right) \mathrm{d}H(\mathbf{z})$$
$$+ M(F_1(x_1), \ldots, F_d(x_d)) + \bar{M}(F_1(x_1), \ldots, F_d(x_d)) - Q'_d(M, C)\Bigg)^2 \mathrm{d}H(\mathbf{x})$$
$$= \int_{[0,1]^d} \Bigg(\sum_{j=1}^{d} \int_{[0,1]^d} \left(\mathbf{1}\left\{v_j \le u_j\right\} - u_j\right)\left(\mathbf{1}\left\{u_j < u_l\ \forall l \ne j\right\} - \mathbf{1}\left\{u_j > u_l\ \forall l \ne j\right\}\right) \mathrm{d}C(\mathbf{u})$$
$$+ M(\mathbf{v}) + \bar{M}(\mathbf{v}) - Q'_d(M, C)\Bigg)^2 \mathrm{d}C(\mathbf{v})$$
$$= \int_{[0,1]^d} \Bigg(\sum_{j=1}^{d} \int_{[0,1]^d} \mathbf{1}\left\{u_j < u_l\ \forall l \ne j\right\}\left(\mathbf{1}\left\{v_j \le u_j\right\} - u_j\right) \mathrm{d}C(\mathbf{u})$$
$$- \sum_{j=1}^{d} \int_{[0,1]^d} \mathbf{1}\left\{u_j > u_l\ \forall l \ne j\right\}\left(\mathbf{1}\left\{v_j \le u_j\right\} - u_j\right) \mathrm{d}C(\mathbf{u})$$
$$+ M(\mathbf{v}) + \bar{M}(\mathbf{v}) - Q'_d(M, C)\Bigg)^2 \mathrm{d}C(\mathbf{v}), \tag{50}$$

we complete the proof. □

Appendix C. Estimation of σ^2

The variance σ^2 in Theorem 1 can be estimated by replacing in the expression of $h(\mathbf{x})$ above, the functions H and F_j by their empirical counter-parts H_n and F_{nj}: Since

$$\widehat{h}(\mathbf{x}) = \sum_{j=1}^{d} \int_{\mathbb{R}^d} \left(\mathbf{1}\left\{x_j \le z_j\right\} - F_j(z_j)\right)$$
$$\cdot \left(\mathbf{1}\left\{F_j(z_j) < F_l(z_l)\ \forall l \ne j\right\} - \mathbf{1}\left\{F_j(z_j) > F_l(z_l)\ \forall l \ne j\right\}\right) \mathrm{d}H_n(\mathbf{z})$$
$$+ M(F_1(x_1), \ldots, F_d(x_d)) + \bar{M}(F_1(x_1), \ldots, F_d(x_d)) + \widehat{Q}'_{nd}(M, C)$$
$$= \frac{1}{n} \sum_{j=1}^{d} \sum_{i=1}^{n} \left(\mathbf{1}\left\{x_j \le X_{ij}\right\} - F_{nj}(X_{ij})\right)$$
$$\cdot \left(\mathbf{1}\left\{F_{nj}(X_{ij}) < F_{nl}(X_{il})\ \forall l \ne j\right\} - \mathbf{1}\left\{F_{nj}(X_{ij}) > F_{nl}(X_{il})\ \forall l \ne j\right\}\right)$$
$$+ M(F_1(x_1), \ldots, F_d(x_d)) + \bar{M}(F_1(x_1), \ldots, F_d(x_d)) + \widehat{Q}'_{nd}(M, C)$$

we have

$$\begin{aligned}\widehat{\sigma}^2 &= \frac{1}{n-1}\sum_{k=1}^{n}\left(\widehat{h}(\mathbf{X}_k) - 2\widehat{Q}'_{nd}(M,C)\right)^2 \\ &= \frac{1}{n-1}\sum_{k=1}^{n}\left(\frac{1}{n}\sum_{j=1}^{d}\sum_{i=1}^{n}\left(\mathbf{1}\left\{X_{kj}\le X_{ij}\right\} - F_{nj}(X_{ij})\right)\right. \\ &\quad\cdot\left(\mathbf{1}\left\{F_{nj}(X_{ij}) < F_{nl}(X_{il})\ \forall l\neq j\right\} - \mathbf{1}\left\{F_{nj}(X_{ij}) > F_{nl}(X_{il})\ \forall l\neq j\right\}\right) \\ &\quad\left. + M(F_1(X_{k1}),\ldots,F_d(X_{kd})) + \bar{M}(F_1(X_{k1}),\ldots,F_d(X_{kd})) - \widehat{Q}'_{nd}(M,C)\right)^2.\end{aligned}$$

Appendix D

Note that according to the calculations preceding (27), for this formula to hold we have to take

$$S_{nd} = \sum_{i_1=1}^{n}\cdots\sum_{i_d=1}^{n}\left(\max(i_1,\ldots,i_d) - \min(i_1,\ldots,i_d)\right).$$

We are going to prove that this is consistent with (28).

Lemma 3. *Let $n, d \in \mathbb{N}$ such that $n \geq 1$ and $d \geq 2$. Then*

$$S_{nd} = (n+1)n^d - 2\sum_{i=1}^{n} i^d = \frac{d-1}{d+1}n^{d+1} - \frac{2}{d+1}\sum_{k=1}^{\lfloor d/2\rfloor}\binom{d+1}{2k}B_{2k}n^{d+1-2k} \tag{51}$$

where $\lfloor x\rfloor$ denotes the floor of x and B_m is the corresponding Bernoulli number, i.e., $\{B_{2k}\}_{k=1}^{\infty} = \{1/6, -1/30, 1/42, -1/30, 5/66, \ldots\}$.

Proof. First, note

$$\begin{aligned}\max(i_1,\ldots,i_d) &= n+1+\max(i_1-n-1,\ldots,i_d-n-1) \\ &= n+1-\min(n+1-i_1,\ldots,n+1-i_d);\end{aligned}$$

Since the terms $n+1-i_r$, with $r = 1,\ldots,d$, take the values $\{1,\ldots,n\}$, we have

$$S_{nd} = (n+1)n^d - 2T_n, \tag{52}$$

where

$$T_n := \sum_{i_1=1}^{n}\cdots\sum_{i_d=1}^{n}\min(i_1,\ldots,i_d).$$

Observe $T_1 = 1$ and $T_n - T_{n-1} = n^d$, thus

$$T_n = \sum_{i=1}^{n} i^d = \frac{n^{d+1}}{d+1} + \frac{n^d}{2} + \sum_{k=2}^{d}\frac{B_k}{k!}\cdot\frac{d!}{(d-k+1)!}\cdot n^{d-k+1}.$$

The second equality is known as Faulhaber's formula and also, very often, Bernoulli's formula (see formula 0.121 in Gradshteyn and Ryzhik [17] and Knuth [20] for more details). Now, putting the expression above back into (52), expression (51) comes out, which completes the proof. □

Appendix E

We next derive a closed-form expression of multivariate Spearman's ρ_d^- in (32) for the FGM d-copula in Example 1. The result follows easily by plugging the formula of the FGM copula function in Equation (8) into the equation (32). Then, simple calculations yield

$$\rho_d^-\left(C_\theta^{\mathbf{FGM}}\right)=\frac{d+1}{2^d-d-1}\left\{2^d\left[\prod_{i=1}^{d}\int_0^1 u_i du_i+\theta\prod_{i=1}^{d}\int_0^1 u_i(1-u_i)du_i\right]-1\right\}=\frac{\theta(d+1)}{3^d(2^d-d-1)}.$$

To obtain the expression of multivariate Spearman's ρ_d^+, we resort to Úbeda-Flores [38, Example 2], who provides an explicit formula of a multivariate version of Spearman's rho, called ρ_d, which is defined as the average of ρ_d^- and ρ_d^+, namely

$$\rho_d\left(C_\theta^{\mathbf{FGM}}\right)=\frac{\rho_d^+\left(C_\theta^{\mathbf{FGM}}\right)+\rho_d^-\left(C_\theta^{\mathbf{FGM}}\right)}{2}=\frac{\theta(d+1)(1+(-1)^d)}{2\cdot 3^d(2^d-d-1)}.$$

Hence, from the two equations above, it is immediate to get the value of ρ_d^+ as

$$\rho_d^+\left(C_\theta^{\mathbf{FGM}}\right)=\frac{\theta(-1)^d(d+1)}{3^d(2^d-d-1)}.$$

Explicit formulae of Kendall's τ_d for the AMH d-copula in Example 3 are not available. However, we will be able to work out a recursive formula to compute $\tau_d'(0)$. To start with, let us denote $\dot{K}_d(t,\theta)=\partial K_d(t,\theta)/\partial\theta$, where $K_d(t,\theta)$ is the Kendall's distribution function of a d-copula C_θ that depends on a single parameter θ. Barbe et al. [2] provide the following expression to compute this function for an Archimedean d-copula with generator ϕ_θ, for $d>2$,

$$K_d(t,\theta)=t+\sum_{i=1}^{d-1}\frac{(-1)^i}{i!}(\phi_\theta(t))^i f_{i-1}(t,\theta), \tag{53}$$

where $f_0(t,\theta)=1/\phi_\theta'(t)$, $\phi_\theta'(t)=\partial\phi_\theta(t)/\partial t$, and for $i\geq 1$, $f_i(t,\theta)=f_0(t,\theta)f_{i-1}'(t,\theta)$, where $f_{i-1}'(t,\theta)=\partial f_{i-1}(t,\theta)/\partial t$. Then, simple calculations based on (53), evaluated at $d=2$, yield

$$\dot{K}_2(t,\theta)=-\frac{1}{\phi_\theta'(t)}\frac{\partial\phi_\theta(t)}{\partial\theta}+\frac{\phi_\theta(t)}{(\phi_\theta'(t))^2}\frac{\phi_\theta'(t)}{\partial\theta}$$

and, for $d>2$, we can define $\dot{K}_d(t,\theta)$ recursively through the following formula,

$$\dot{K}_d(t,\theta)=\dot{K}_{d-1}(t,\theta)+\frac{(-1)^{d-1}}{(d-1)!}\frac{\partial}{\partial\theta}\left[(\phi_\theta(t))^{d-1}f_{d-2}(t,\theta)\right]. \tag{54}$$

Now, taking into account (31), one easily deduces

$$\tau_d'(\theta)=\frac{\partial\tau_d(\theta)}{\partial\theta}=-\frac{2^d}{2^{d-1}-1}\int_0^1\dot{K}_d(t,\theta)\,\mathrm{d}t, \tag{55}$$

(see Genest et al. [16]) and putting (54) back into (55), a direct evaluation of $\tau_d'(\theta)$ at $\theta=0$ gives the following recursion for any Archimedean d-copula with generator ϕ_θ, for $d>2$,

$$\tau_d'(0)=\frac{2^{d-1}-2}{2^{d-1}-1}\tau_{d-1}'(0)-\frac{2^d}{2^{d-1}-1}\frac{(-1)^{d-1}}{(d-1)!}\int_0^1\frac{\partial}{\partial\theta}\left[(\phi_\theta(t))^{d-1}f_{d-2}(t,\theta)\right]\Big|_{\theta=0}\mathrm{d}t.$$

In particular, for the AMH d-copula in Example 3, tedious calculations yield

$$\tau_2'\left(C_0^{\mathbf{AMH}}\right)=\tau_3'\left(C_0^{\mathbf{AMH}}\right)=0.222,\ \ \tau_4'\left(C_0^{\mathbf{AMH}}\right)=0.194,\ \ \tau_5'\left(C_0^{\mathbf{AMH}}\right)=0.156,$$

and so on.

References

[1] M.M. Ali, N.N. Mikhail, M.S. Haq, A class of bivariate distributions including the bivariate logistic, J. Multivar. Anal. 8 (1978) 405–412.
[2] P. Barbe, C. Genest, K. Ghoudi, B. Rémillard, On Kendall's process, J. Multivar. Anal. 58 (1996) 197–229.
[3] G. Beliakov, E. de Amo, J. Fernández-Sánchez, M. Úbeda-Flores, Best-possible bounds on the set of copulas with a given value of Spearman's footrule, Fuzzy Sets Syst. 428 (2022) 138–152; Corrigendum 428, 153–155.
[4] D.G. Clayton, A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence, Biometrika 65 (1978) 141–151.
[5] C.M. Cuadras, J. Augé, A continuous general multivariate distribution and its properties, Commun. Stat., Theory Methods 10 (1981) 339–353.
[6] K. Decancq, Measuring cumulative deprivation and affluence based on the diagonal dependence diagram, Metron 78 (2020) 103–117.
[7] A. Dolati, M. Úbeda-Flores, On measures of multivariate concordance, J. Probab. Stat. Sci. 4 (2) (2006) 147–163.
[8] F. Durante, C. Sempi, Principles of Copula Theory, Chapman & Hall/CRC, Boca Raton, 2016.
[9] F. Durante, R.B. Nelsen, J.J. Quesada-Molina, M. Úbeda-Flores, Pairwise and global dependence in trivariate copula models, in: A. Laurent, O. Strauss, B. Bouchon-Meunier, R.R. Yager (Eds.), Information Processing and Management of Uncertainty in Knowledge-Based Systems, Springer International Publishing, Montpellier, France, 2014, pp. 243–251.
[10] S. Fuchs, Y. McCord, On the lower bound of Spearman's footrule, Depend. Model. 7 (2019) 126–132.
[11] J.E. García, V. González-López, R.B. Nelsen, A new index to measure positive dependence in trivariate distributions, J. Multivar. Anal. 115 (2013) 481–495.
[12] C. García-Gómez, A. Pérez, M. Prieto-Alaiz, Copula-based analysis of multivariate dependence patterns between dimensions of poverty in Europe, Rev. Income Wealth 67 (1) (2021) 165–195.
[13] C. Genest, P. Jaworski, On the class of bivariate Archimax copulas under constraints, Fuzzy Sets Syst. 415 (2021) 37–53.
[14] C. Genest, J. Nešlehová, N.B. Ghorbal, Spearman's footrule and Gini's gamma: a review with complements, J. Nonparametr. Stat. 22 (2010) 937–954.
[15] C. Genest, J. Nešlehová, N.B. Ghorbal, Estimators based on Kendall's tau in multivariate copula models, Aust. N. Z. J. Stat. 53 (2) (2011) 157–177.
[16] C. Genest, J-F. Quessy, B. Rémillard, Goodness-of-fit procedures for copula models based on the probability integral transformation, Scand. J. Stat. 33 (2) (2006) 337–366.
[17] I.S. Gradshteyn, I.M. Ryzhik, Table of Integrals, Series and Products, 7th ed., Elsevier/Academic Press, Amsterdam, 2007.
[18] M. Hofert, I. Kojadinovic, M. Maechler, J. Yan, Multivariate dependence with copulas, R package version 1.0-1, https://CRAN.R-project.org/package=copula, 2020.
[19] H. Joe, Multivariate concordance, J. Multivar. Anal. 35 (1990) 12–30.
[20] D.E. Knuth, Johann Faulhaber and sums of powers, Math. Comput. 61 (203) (1993) 277–294.
[21] E. Liebscher, On a multivariate version of Spearman's correlation coefficient for regression: properties and applications, Asian J. Stat. Sci. 1 (2) (2021) 123–150.
[22] A.J. McNeil, J. Nešlehová, Multivariate Archimedean copulas, d-monotone functions and ℓ_1-norm symmetric distributions, Ann. Stat. 37 (5B) (2009) 3059–3097.
[23] M. Mesfioui, J.-F. Quessy, Concordance measures for multivariate non-continuous random vectors, J. Multivar. Anal. 101 (10) (2010) 2398–2410.
[24] R.B. Nelsen, Nonparametric measures of multivariate association, in: L. Rüschendorf, B. Schweizer, M.D. Taylor (Eds.), Distributions with Given Marginals and Related Topics, Institute of Mathematical Statistics, Hayward, CA, 1996, pp. 223–232.
[25] R.B. Nelsen, Concordance and Gini's measure of association, J. Nonparametr. Stat. 9 (1998) 227–238.
[26] R.B. Nelsen, Concordance and copulas: a survey, in: C.M. Cuadras, J. Fortiana, J.A. Rodriguez-Lallena (Eds.), Distributions with Given Marginals and Statistical Modelling, Kluwer, Dordrecht, 2002, pp. 169–178.
[27] R.B. Nelsen, An Introduction to Copulas, 2nd ed., Springer, New York, 2006.
[28] R.B. Nelsen, M. Úbeda-Flores, Directional dependence in multivariate distributions, Ann. Inst. Stat. Math. 64 (2012) 677–685.
[29] A. Pérez, M. Prieto-Alaiz, A note on nonparametric estimation of copula-based multivariate extensions of Spearman's rho, Stat. Probab. Lett. 112 (2016) 41–50.
[30] A. Pérez, M. Prieto-Alaiz, Measuring the dependence among dimensions of welfare: a study based on Spearman's footrule and Gini's gamma, Int. J. Uncertain. Fuzziness Knowl.-Based Syst. 24 (1) (2016) 87–105.
[31] J.-F. Quessy, Theoretical efficiency comparisons of independence tests based on multivariate versions of Spearman's rho, Metrika 70 (2009) 315–338.
[32] F. Schmid, R. Schmidt, Multivariate extensions of Spearman's rho and related statistics, Stat. Probab. Lett. 77 (2007) 407–416.
[33] J.B. Seaborn, Hypergeometric Functions and Their Applications, Springer-Verlag, New York, 1991.
[34] R.J. Serfling, Approximation Theorems of Mathematical Statistics, Wiley, New York, 1980.
[35] A. Sklar, Fonctions de répartition à n dimensions et leurs marges, Publ. Inst. Stat. Univ. Paris 8 (1959) 229–231.
[36] C. Spearman, Footrule for measuring correlation, Br. J. Psychol. 1 (2) (1906) 89–108.
[37] N.A. Stepanova, Multivariate rank tests for independence and their asymptotic efficiency, Math. Methods Stat. 12 (2) (2003) 197–217.
[38] M. Úbeda-Flores, Multivariate versions of Bloomqvist's beta and Spearman's footrule, Ann. Inst. Stat. Math. 57 (2005) 781–788.
[39] A.W. van der Vaart, Asymptotic Statistics, Cambridge University Press, Cambridge, 1998.
[40] E.F. Wolff, N-dimensional measures of dependence, Stochastica 4 (1980) 175–188.