An AUK-based index for measuring and testing the joint dependence of a random vector

Georgios Afendras*
Department of Mathematics, Aristotle University of Thessaloniki, Thessaloniki, Greece

Marianthi Markatou
Department of Biostatistics and Jacobs School of Medicine and Biomedical Sciences, The State University of New York at Buffalo, Buffalo, NY, USA

Albert Vexler
Department of Biostatistics, The State University of New York at Buffalo, Buffalo, NY, USA

Summary. We present an index of dependence that allows one to measure the joint or mutual dependence of a $d$-dimensional random vector with $d > 2$. The index is based on a $d$-dimensional Kendall process. We further propose a standardized version of our index of dependence that is easy to interpret, and provide an algorithm for its computation. We discuss tests of total independence based on consistent estimates of the area under the Kendall curve. We evaluate the performance of our procedures via simulation, and apply our methods to a real data set.

MSC 2010 subject classifications: Primary 62H20, 62H05, 62E10; secondary 62-09, 62G99.

Key words and phrases: Copulas; Dependence axioms; Index of dependence; Kendall-function/plot; Measure of association/dependence; Structure theory; Testing complete independence.

1. Introduction and Literature Review

Advances in technology allow the generation and collection of very large data sets, where the dependencies among the different variables that comprise these data sets are often of considerable interest. In this paper, we consider ways to measure joint or mutual dependence among the components of a $d$-dimensional random vector and to construct test procedures for mutual or total independence.

There is substantial literature on the construction of indices of dependence between two random variables. These include the relevant concept of correlation, for example Pearson (1895) and Spearman’s rho (1904). More recently, Székely et al. (2007) and Székely and Rizzo (2009, 2014) proposed the concept of distance correlation, that represents one of the most important breakthroughs in measuring dependence between two random vectors. Furthermore, a generalization of the concept of distance correlation was recently proposed by Shen et al. (2020). Other proposals include a projection correlation (Zhu et al., 2017), Bergsma and Dassios (2014) and Genest et al. (2019).

We propose an index of dependence that is based on the area under the Kendall curve, generalizing work by Vexler et al. (2019) from the two-dimensional to the $d$-dimensional case.

*Corresponding author
E-mail addresses: gafendra@math.auth.gr (Georgios Afendras); markatou@buffalo.edu (Marianthi Markatou); avexler@buffalo.edu (Albert Vexler)
The proposed index measures the dependence between the \( d \) components of a random vector. We then construct a test of total independence based on the estimated area under the Kendall curve. This task is different from the task of testing pairwise independence in a collection of random variables.

The literature in statistics includes a few nonparametric procedures to test mutual or joint dependence. These include Simon (1977) who proposed a nonparametric test based on Kendall’s tau. Other tests include Kankainen (1995) and Karvanen (2005). The later author developed a restamping test for testing total independence that is based on the test proposed by Kankainen, for the case of stationary time series.

Pfister et al. (2018) investigated also the problem of testing mutual or joint independence and proposed tests based on the \( d \)-variable Hilbert-Schmidt independence criterion, dHSIC, thereby extending Kankainen (1995) to the flexible framework of kernel methods. These authors developed their test based on the idea of probability embeddings of marginals and the joint distribution of the random vector \( X \).

Our contribution is two-fold. First, we propose an algorithm to construct an index of dependence that is based on the Euclidean distance between the vector of appropriately defined areas under the Kendall curve and the \( d \)-dimensional vector with components equal to 1/2. We also present the standardized version of our index of dependence that is easily interpretable, and we note that these indices are copula-based.

Next, we propose a practical, fast to compute test of mutual independence that is also based on the estimate of the area under the Kendall curve. We discuss the asymptotic distribution of our test statistic under the null hypothesis of mutual independence and investigate the performance characteristics of our test via simulation. Furthermore, we provide comparisons with the dHSIC test. The results of our study show that the proposed testing procedure is practical and easily computable — the computation time required to carry out our test is considerably less than the time required to compute the corresponding dHSIC test. Furthermore, the newly proposed test outperforms in terms of power the dHSIC test in smaller samples and performs as the dHSIC test in large samples. The power results apply to all distributions from where samples were generated with the exception of the distributions \( U\{C(0, 1)\} \) and spiral 3 and 4 distributions.

The paper is organized as follows. Section 2 offers motivation for constructing the index of dependence and presents discussion on appropriate axioms that a measure of dependence must satisfy. Section 3 briefly presents the bivariate case, while Section 4 extends the index to the case where the dimension \( d \) is greater than 2. This section also offers an algorithm for the computation of a standardized version of our index. Section 5 proves that the proposed index satisfies the axioms a function must satisfy in order to be a measure of dependence, while Section 6 discuss the estimation of the area under the Kendall curve and the two indices of dependence. Section 7 introduces the test of mutual independence, while Sections 8 and 9 present simulation results and a real data example. Section 10 concludes with a short summary and discussion of the results. Proofs of all relevant theorems are given in the appendix.

### 2. Motivation, Notation and Dependence Axioms

Let \( X_1, \ldots, X_d \) be a set of \( d \) unidimensional random variables (rvs). Let \( d_1 \) and \( d_2 \) be two positive integers such that \( d_1 + d_2 = d \); consider a partitioning of \( X \)'s into two sets of dimensions \( d_1 \) and \( d_2 \), say \( X_1, \ldots, X_{d_1} \) and \( X_{d_1+1}, \ldots, X_d \). Consider the random vectors \( X_1 = (X_1, \ldots, X_{d_1})^T, X_2 = (X_{d_1+1}, \ldots, X_d)^T \).
and Székely (2019) propose four natural axioms for dependence measures, see Axioms
measure on by Schweizer and Wolff (1981), see Axioms
general, refer to the statement in
Hofert, Oldford, Prasad and Zhu (2019) establish axioms for dependence measures, that also, in
and Székely only refer to the case of the statement in
suitable for
Specifically, for the cases where
\( R \)
\( X \)
\( x \)
\( d \)
\( \rho \)
\( \Sigma \)

Suppose \( X \in \mathcal{X}_0^d \) and assume that we are interested in measuring the following two kinds of
dependence, measured as distances between appropriate cdfs:
\( \mathcal{D}_1 \): measuring the distance between \( F(x) \) and \( F_1(x_1)F_2(x_2) \);
\( \mathcal{D}_2 \): measuring the distance between \( F(x) \) and \( \prod_{j=1}^d F_j(x_j) \).
In other words, in \( \mathcal{D}_1 \) we are interested in measuring the pairwise-dependence between \( X_1 \) and
\( X_2 \); while in \( \mathcal{D}_2 \) we are interested in measuring the dependence structure of the joint distribution
function of \( X \). Thus, in \( \mathcal{D}_2 \) we aim to measure joint or mutual dependence. In the bivariate case,
i.e. \( d = 2 \) and so \( d_1 = d_2 = 1 \), the statements \( \mathcal{D}_1 \) and \( \mathcal{D}_2 \) are identical; while these differ from
one another when \( d > 2 \).

We focus in the multivariate case where \( d > 2 \). Consider the rvs \( U_j, j = 1, \ldots, d \), independent
and identically distributed from \( U(0, 1) \) such that \( X \) and \( U_j \)'s are independent, and define the rvs
\( T \equiv F(X) \); \( \Pi \equiv U_1 \cdots U_d \); \( E \equiv - \ln \Pi \).

\( \Sigma_d(\rho) \)

Let the components of a random vector \( X \) be standardized (each marginal distribution has zero
mean and variance one), and denote its variance/covariance matrix by \( \Sigma_d(\rho) \) with
\( \rho = (\rho_{ij}) \in \mathbb{R}^{n(n-1)/2}, 1 \leq i < j \leq d \), that is
\( \Sigma_d(\rho) = (\sigma_{ij})_{d \times d} \) with \( \sigma_{ii} = 1 \) and \( \sigma_{ij} = \sigma_{ji} = \rho_{ij} \) for \( i < j \).

\( 2.2a \)

Specifically, for the cases where \( \rho_{ij} = \rho \) for all \( i < j \), we denote the variance/covariance matrix
by \( \Sigma_d(\rho) \), that is,
\( \Sigma_d(\rho) = (\sigma_{ij})_{d \times d} \) with \( \sigma_{ii} = 1 \) and \( \sigma_{ij} = \rho \) for \( i \neq j \).

\( 2.2b \)

In 1959, Rényi’s fundamental article defined a set of axioms that a measure of dependence
for a pair of random variables must satisfy; these axioms have been studied and suitably modified
by Schweizer and Wolff (1981), see Axioms \( A^R_1 \) to \( A^R_4 \) in Appendix A. In a recent article, Móri
and Székely (2019) propose four natural axioms for dependence measures, see Axioms \( A^MS_1 \)
to \( A^MS_4 \) in Appendix A. Rényi’s axioms refer only to the 2-dimensional case; while those of Móri
and Székely only refer to the case of the statement in \( \mathcal{D}_1 \). Similar to Móri and Székely (2019),
Hofert, Oldford, Prasad and Zhu (2019) establish axioms for dependence measures, that also, in
general, refer to the statement in \( \mathcal{D}_1 \). We discuss now a modification of Rényi’s axioms that is
suitable for \( \mathcal{D}_2 \).

Let \( X^d \subset \mathcal{X}_0^d \). According to Rényi, a mapping \( \delta: X^d \rightarrow [0, 1] \) is called a mutual dependence
measure on \( X^d \) if
Remark above analysis, there are positive values of 
Consequently, 
\[
\det \Sigma \text{ is a positive definite matrix and so its determinant is non-negative, it follows that } \\
\text{while } \frac{1}{\rho}^2 \frac{2}{1 + 2 \rho} \text{ since a variance/covariance matrix is a positive definite matrix and so its determinant is non-negative, it follows that } \\
\rho \geq -1/2. \\
\text{Consequently, } \Sigma_3(\rho) \text{ is a variance/covariance matrix if and only if } -1/2 \leq \rho \leq 1. \\
\text{Based on the above analysis, there are positive values of } \rho \text{ for which the fact that } \\
\Sigma_3(\rho) \text{ can be assumed as a variance/covariance matrix of a } n \text{-dimensional random vector does not imply that } \\
\Sigma_3(-\rho) \text{ can also be assumed as a variance/covariance matrix. For example, set } \rho = 0.7; \text{ then, } \\
\Sigma_3(0.7) \text{ is a positive definite } 3 \times 3 \text{ matrix and hence we can consider the normal distribution } \\
N_d(\mathbf{0}, \Sigma_d(0.7)), \text{ while } \Sigma_3(-0.7) \text{ is not a positive definite matrix and so the notation } \\
N_d(\mathbf{0}, \Sigma_d(-0.7)) \text{ makes no sense.} \\
\text{In general, the determinant of } \Sigma_d(\rho), \text{ given by } (2.2b), \text{ is } \\
\det \{ \Sigma_d(\rho) \} = (1 - \rho)^{d-1} \{1 + \rho(d - 1)\}. \\
\text{Therefore, this matrix is a variance/covariance matrix if and only if } (1 - d)^{-1} \leq \rho \leq 1. \text{ Let } \\
\mathbf{X} \sim N_d(\mathbf{0}, \Sigma_d(\rho)); \text{ then, the acceptable values of } \rho \text{ are } \\
(1 - d)^{-1} \leq \rho \leq 1. \text{ If Axiom } A^B_i \text{ can be modified for general } d, \text{ then a measure of dependence will take the same value for both } \\
N_d(\mathbf{0}, \Sigma_d(\rho)) \text{ and } N_d(\mathbf{0}, \Sigma_d(-\rho)) \text{ distributions for all } \rho; \text{ on the other hand, for } d > 2 \text{ and each } \\
(d - 1)^{-1} < \rho \leq 1 \text{ the notation } N_d(\mathbf{0}, \Sigma_d(\rho)) \text{ denotes the } d \text{-dimensional normal distribution with mean vector } \mathbf{0} \text{ and variance/covariance matrix } \\
\Sigma_d(\rho), \text{ while the notation } N_d(\mathbf{0}, \Sigma_d(-\rho)) \text{ does not correspond to any distribution. } \\
\text{(b) One may incorrectly suggest an extension of } A^B_i \text{ as follows. For a } d \text{-dimensional normal distribution for fixed } \rho_{ij} \text{ with } (i, j) \neq (i_0, j_0), \text{ a measure of mutual dependence must be an increasing function of } |\rho_{i_0, j_0}|. \text{ To explain why the aforementioned statement is incorrect, we consider the } 3 \text{-dimensional normal distributions } \\
F \sim N_3(\mathbf{0}, \Sigma_F) \text{ and } G \sim N_3(\mathbf{0}, \Sigma_G), \text{ where the variance/covariance matrices are } \\
\Sigma_F = \begin{pmatrix} 1 & -0.5 & 0.5 \\ -0.5 & 1 & -0.5 \\ 0.5 & -0.5 & 1 \end{pmatrix} \text{ and } \Sigma_G = \begin{pmatrix} 1 & -0.5 & -0.5 \\ -0.5 & 1 & -0.5 \\ -0.5 & -0.5 & 1 \end{pmatrix}. \text{ Suppose } \mathbf{X} \sim G. \\
\text{Because of } \det(\Sigma_G) = 0, \text{ there is constant vector } \mathbf{c} \in \mathbb{R}^3 \text{ such that } \mathbf{c}^T \mathbf{X} \text{ is a constant with } \\
\text{probability one, that is, } \mathbf{X} \text{ lies on a subset of } \mathbb{R}^3 \text{ of dimension } 2. \text{ The support of } F \text{ is } S_F = \mathbb{R}^3 \text{ while the support of } G \text{ is } S_G \subseteq \mathbb{R}^3 \text{ with } \dim(S_G) = 2. \text{ Figures } 8.1a \text{ and } 8.1b \text{ present scatterplots based on random samples of size } 1000 \text{ from } F \text{ and } G \text{ respectively. } \\
\text{It is clear that } G \text{ has a stronger dependence structure than } F, \text{ and so it is an abnormal situation a measure of dependence to have } \\
\text{the same value for } F \text{ and } G. \\
\text{(c) In view of (a) and (b), such modification of Axiom } A^B_i \text{ makes no sense for dimension } d > 2.
On the other hand, according to the definition by Móri and Székely, a mapping \( \delta : \mathcal{X}^d \to [0, 1] \) is called a structure dependence measure on \( \mathcal{X}^d \) if

- \( A_1' : \delta(X) = 0 \) if and only if \( X_1, \ldots, X_n \) are independent rvs.
- \( A_2' : \delta(X) \) is invariant with respect to all similarity transformations of \( \mathbb{R}^d \); that is, \( \delta(LX) = \delta(X) \) where \( L \) is similarity transformation of \( \mathbb{R}^d \).
- \( A_3' : \delta(X) = 1 \) if and only if there is a similarity transformation \( L \) such that \( LX = (X_L, \ldots, X_L)^T \in \mathcal{X}^d \) with probability 1.
- \( A_4' : \delta(\cdot) \) is continuous; that is, if \( X \) and \( X_n, n = 1, 2, \ldots \), belong to \( \mathcal{X}^d \) such that \( X_n \) is uniform integrable and \( X_n \xrightarrow{a.s.} X \), then \( \delta(X_n) \xrightarrow{} \delta(X) \).

We note that Móri and Székely (2019) as well as Hofert et al. (2019) avoid to establish an axiom analogous to \( A_4' \). In Section 5 we prove that our proposed index satisfies both sets of axioms, the modified Rényi (\( A_1' - A_6' \)) and Móri and Székely (\( A_1' - A_4' \)).

### 3. The bivariate case

We aim to construct a measure of mutual dependence that satisfies axioms \( A_1 - A_8 \) discussed in the previous section. For this purpose, in this section we briefly discuss the bivariate case that has been studied by Vexler et al. (2019).

Consider the rvs \( T \) and \( \Pi \) defined by (2.1). The Kendall distribution function is defined as \( K(t) = \mathbb{P}(T \leq t) \). Since the distribution function of \( \Pi \) that corresponds to the independent case is \( K_{\Pi}(t) = t - t \ln t, t \in [0, 1] \), the Kendall-plot is the visualization of the curve \((K(t), t - t \ln t), t \in [0, 1]\), and compared to the diagonal line \((t - t \ln t, t - t \ln t), t \in [0, 1]\), it provides a rich source of information reference the association between the components of \( X \). In the light of an ROC analysis, Vexler et al. (2019) show that the area under the Kendall curve is \( \text{AUK} = -\int_0^1 K(t) \ln t \, dt \), and then that

\[
\text{AUK} = \mathbb{E}(1 - T + T \ln T),
\]

which is useful both, for direct evaluation and for simulation studies based on the AUK; it is also used for nonparametric estimation of AUK based on a random sample of pairs from \( F \).

Vexler et al. (2019) offer motivation for the use of an extension of the Kendall plot to a multi-panel Kendall plot that consists of the ordinary Kendall plots of the four orthogonal rotations of \( X \); in particular, these are the Kendall plots that correspond to the Kendall functions \( K_i(t) = \mathbb{P}(H_i(X) \leq t), i = 1, \ldots, 4 \), where

\[
\begin{align*}
H_1(x_1, x_2) &= \mathbb{P}(X_1 \leq x_1, X_2 \leq x_2), & H_2(x_1, x_2) &= \mathbb{P}(X_1 > x_1, X_2 \leq x_2), \\
H_3(x_1, x_2) &= \mathbb{P}(X_1 \leq x_1, X_2 > x_2), & H_4(x_1, x_2) &= \mathbb{P}(X_1 > x_1, X_2 > x_2).
\end{align*}
\]

Using these Kendall functions, they provide the four corresponding AUKs \( \text{AUK}_i, i = 1, \ldots, 4 \).

Based on \( D = (\text{AUK}_1, \ldots, \text{AUK}_4)^T \) they establish an index of dependence between \( X_1 \) and \( X_2 \) by the relation

\[
I = (8/5)^{1/2} \| D - A \|,
\]

where \( A = (1/2, \ldots, 1/2)^T \in \mathbb{R}^d \) and \( \|x\| \) denotes the Euclidean norm of the vector \( x \). They also present a standardized index of dependence based on \( I \) that ensures a linear mapping with \(|\rho| \) in
the case of the bivariate normal distribution, and they give an approximation of this standardized
index by

\[ I^* = \varphi_2(I), \]

where \( \varphi_2(t) = 2.070t + 0.061t^2 - 2.471t^3 + 1.307t^4 + 0.033t^5, t \in [0, 1]. \)

In what follows, we will extend the index of dependence discussed above to the \( d \)-dimensional
\( \mathbf{X} \) be a \( d \)-dimensional random vector which has continuous components. For the dimension
\( I^* = \varphi_2(I), \)

\( d \) of \( \mathbf{X}, \) consider the rvs \( T, \Pi \) and \( E \) defined by (2.1). A natural extension from the bivariate
case to \( d \)-dimensional case of the Kendall distribution function is given by \( K(t) = \mathbb{P}(T \leq t). \)

\( \) When we are interested in measuring the dependence structure of \( \mathbf{X}, \) according to \( D_2, \) we can plot this Kendall function against the cdf of \( \mathbf{X}, \)

\( \) which corresponds to the case of independence of the components of \( \mathbf{X}; \) and then, we compare this curve with the diagonal line in \([0, 1] \times [0, 1]. \)

\( \) Namely, we compare the curve \( (K(t), K_{II}(t)) \) with the diagonal line \( (K_{II}(t), K_{II}(t)), t \in [0, 1]. \)

\( \) Both Kendall cdf, \( K_{II} (\cdot), \) and probability density function (pdf), \( k_{II} (\cdot), \) of \( \Pi \) may be computed

\( \) in close forms. Specifically, as is evident \( E \sim \text{Erlang}(d, 1), \)

\( \)

\( K_{II}(t) = t \sum_{j=0}^{d-1} \frac{(-1)^k}{k!} \ln^k t \) and \( k_{II}(t) = \frac{(-1)^{d-1}}{(d-1)!} \ln^{d-1} t, \quad 0 < t < 1. \)

\( \) (4.1)

\( \)

\( \) Observe now that \( 2E \) follows the chi-square distribution with \( 2d \) degrees of freedom, denoted as

\( \) \( E \sim \chi^2_{2d}. \) By definition of \( E, \) see (2.1), the cdf and pdf of \( \Pi \) can be presented through the cdf and

\( \) pdf of \( \chi^2_{2d}; \) specifically,

\( \)

\( K_{II}(t) = 1 - F_{\chi^2_{2d}} (-2 \ln t) \) and \( k_{II}(t) = \frac{2}{t} f_{\chi^2_{2d}} (-2 \ln t), \quad t > 0. \)

\( \) The above formulas are appropriate for computation.

\( \)

\( \) The area under the Kendall (AUK) curve is \( \int_0^1 K(t) dK_{II}(t) = \int_0^1 K(t) k_{II}(t) dt, \) that is

\( \)

\( \text{AUK} = \mathbb{P}(T \leq \Pi). \)

\( \)

\( \) Since \( \mathbb{P}(T \leq \Pi) = \mathbb{E}\{\mathbb{P}(T \leq \Pi | T)\} = \mathbb{E}\{\mathbb{P}(T \leq \Pi | T)\} = \mathbb{E}\{1 - \mathbb{P}(\Pi < T | T)\}, \)

\( \)

\( \text{AUK} = \mathbb{E}\{K(\Pi)\} = \mathbb{E}\{1 - K_{II}(T)\}. \)

\( \)

\( \) Note that the formulas of \( K_{II} \) and AUK in bivariate case follow from the above corresponding

\( \) formulas of the general case by setting \( d = 2. \)

\( \)

\( \) The distance between AUK and \( 1/2 \) may itself be used as a measure of divergence between

\( \) \( F \) and \( \prod_{j=1}^{d} F_j \). But, it cannot be used as a measure of dependence because it does not satisfy the

\( \) axioms of a measure of dependence that are listed in Section 2, Axioms \( A_1 \) to \( A_8 \) or \( A'_1 \) to \( A'_4 \).

\( \)

\( \) We now investigate the four Kendall functions that lead to this multi-panel plot in the bivariate

\( \) case. This investigation offers understanding that is used in the extension of the work for the case of

\( d > 2. \) Observe that, since \( X_1 \) and \( X_2 \) are continuous, the functions \( H_1, \ldots, H_4 \) that are defined by

\( \)
(3.1) are the cdfs of \( X_1 = (X_1, X_2)^T, X_2 = (-X_1, X_2)^T, X_3 = (-X_1, -X_2)^T \) and \( X_4 = (X_1, -X_2)^T \) respectively; hence, the Kendall distribution functions \( K_1, \ldots, K_4 \) are the corresponding ordinary Kendall distribution functions of \( X_1, \ldots, X_4 \). Therefore, the four rotations of a bivariate random vector are leading to the construction of the corresponding Kendall functions, and so to the construction of the AUK-vector, \( D = (\text{AUK}_1, \ldots, \text{AUK}_4)^T \), that can be presented in an algebraic way, through all possible multiplications of the components \( X_1 \) and \( X_2 \) with \( \pm \). This representation provides us a natural algebraic manner to extend the AUK-method in highest dimensional random vectors.

In view of the above analysis, we consider the set of the orthogonal “rotations” of \( X \), say \( \mathcal{R}(X) \). Analytically, define the subset \( S^d \) of \( \mathbb{R}^d \) that contains all possible vectors with elements \( \pm 1 \), that is
\[
S^d = \{ -1, +1 \}^d = \{ s_j \}_{j=1}^{2^d}.
\]
Then, the set of the orthogonal “rotations” of \( X \) is
\[
\mathcal{R}(X) = \{ X_j = \text{diag}(s_j)X : s \in S^d \},
\]
where \( \text{diag}(s_j) \) denotes the \( d \times d \) diagonal matrix with diagonal elements the elements of \( s_j \). For \( X_j \in \mathcal{R}(X) \), \( j = 1, \ldots, 2^d \), define
\[
T_j = F_{X_j}(X_j) \quad \text{and} \quad K_j(t) = \mathbb{P}(T_j \leq t).
\]
The corresponding, under the Kendall curve, areas are
\[
\text{AUK}_j = \mathbb{P}(T_j \leq \Pi) = \mathbb{E}\{ 1 - K_{\Pi}(T_j) \}.
\]
Define \( D = (\text{AUK}_1, \ldots, \text{AUK}_{2^d})^T \) and \( A = (1/2, \ldots, 1/2)^T \in \mathbb{R}^{2^d} \).

Let now \( X \) be a continuous rv, and consider the random vectors \( X_1 = (X, \ldots, X)^T \) and \( X_2 = (X, \ldots, -X)^T \) in \( \mathbb{R}^d \). Write \( X_2 = f(X_1) \), where \( f = (f_1, \ldots, f_d)^T \) with \( f_i(x) = x \) for \( i = 1, \ldots, d - 1 \) and \( f_d(x) = -x \). Obviously, \( f_i \) is a strictly monotone function for all \( i = 1, \ldots, d \), and \( f \) is a similarity transformation; so, for any measure of dependence \( \delta \) it is required \( \delta(X_1) = \delta(X_2) \). On the other hand, the Kendall functions of \( X_1 \) and \( X_2 \) are \( K_1(t) = K_{X_1}(t) = t \) and \( K_2(t) = K_{X_2}(t) = 1, 0 \leq t \leq 1 \). By application of the formula \( \text{AUK} = \int_0^1 K(t)k_{\Pi}(t) dt \), we get \( \text{AUK}_1 = 2^{-d} \) and \( \text{AUK}_2 = 1 \) (for computational details see in the proof of (4.2) in Appendix D), and hence, the distances between \( \text{AUK}_1 \) and \( 1/2 \) (the value that corresponds to the independence case), \( i = 1, 2, \) are \( 1/2 - 2^{-d} \) and \( 1/2 \) respectively. Because of Axiom \( A_6 \) or Axiom \( A_2' \), the AUK of \( X \in \mathcal{X}^d \) cannot be used for measuring the dependence structure of \( X \).

In general, due to \( A_6 \) or \( A_2' \), \( \delta(X_j) = \delta(X) \) for all \( X_j \in \mathcal{R}(X) \). Hence, we propose an AUK-based index of dependence of the form \( c_d \| D - A \| \), where the constant \( c_d \) is chosen such that the index takes the value 1 whenever \( X \) is of the form \( (X, \ldots, X)^T \) for some continuous rv \( X \). Let \( X \) be a continuous rv and set \( X = (X, \ldots, X)^T \in \mathcal{X}^d \). Then, the set of random vectors \( \mathcal{R}(X) \) contains exactly two random vectors with all components being the same, \( (X, \ldots, X)^T \) and \( (-X, \ldots, -X)^T \). Each of \( 2^d - 2 \) remaining random vectors has components that are rv \( X \) and \( -X \). Each of \( (X, \ldots, X)^T \) and \( (-X, \ldots, -X)^T \) has the same Kendall function \( t, 0 \leq t \leq 1 \), while each of \( 2^d - 2 \) remaining random vectors has the same Kendall function \( 1, 0 \leq t \leq 1 \); therefore, the constant \( c_d \) is (for computational details see in the proof of (4.2) in Appendix D)
\[
c_d = \left( 2^{d-2} - 2^{1-d} + 2^{1-2d} \right)^{-1/2}.
\]
Thus, the proposed index of dependence takes the form

\[
I = \left(2^{d-2} - 2^{1-d} + 2^{1-2d}\right)^{-1/2} \|D - A\|. \tag{4.3}
\]

Similar to the bivariate case, we are interested in finding the easily interpreted standardized index, \(I_{st}\). This index is a function of \(I\) that maps \(I \mapsto \rho_j\) whenever \(X \sim N_d(\mathbf{0}, \Sigma_d(\rho))\), \(0 \leq \rho \leq 1\); namely, \(I_{st} = \phi_d(I)\). The index does not have a closed form expression even in the bivariate case, but it can be approximated by \(I = \phi_d(I)\).

For the trivariate normal distribution, \(X \sim N_3(\mathbf{0}, \Sigma_3(\rho))\), we compute the true value of \(I\) for various values of \(\rho\), see Table III.1 in Supplementary Material. In similar way as in Vexler et al. (2019), i.e. by applying the Lagrange interpolation polynomial formula, a polynomial approximation of \(\phi_3\) is

\[
\phi_3(t) = 1.62t + 4.45t^2 - 13.48t^3 + 12.13t^4 - 3.72t^5, \quad t \in [0, 1].
\]

The preceding approximation function \(\phi_3: [0, 1] \to [0, 1]\) is a strictly increasing function. Moreover, \(\phi_3\) is a perfect fit to the true \(\phi_3\), see Figure 4.1a; therefore, the approximate standardized index \(I^*\) has an almost perfect fit to the standardized index \(I_{st}\), that is, it has an almost perfect fit to the identity function of \(\rho\) when the underlying distribution is \(N_3(\mathbf{0}, \Sigma_3(\rho))\) with \(\rho \in [0, 1]\), see Figure 4.1b.

Both \(\phi_2: [0, 1] \to [0, 1]\) and \(\phi_3: [0, 1] \to [0, 1]\) are strictly increasing functions. The Lagrange interpolation polynomial formula, that has been used for the cases of dimension 2 and 3, often leads to find interpolating polynomials that are not strictly monotone; hence, it is wise to avoid for dimensions greater than 3. For any specific value of \(d\), we propose the use the \texttt{MonoPoly} R package (Turlach and Murray, 2019) to find an appropriate, strictly monotone, polynomial for approximating the standardized index. The procedure is presented by Algorithm 4.1: in Section V of the Supplementary Material, we provide a realization of Step 2 of the algorithm.
Algorithm 4.1 Computation of the approximate standardized index $I^*$.

1: Let $0 \leq \rho \leq 1$ and consider $X_\rho \sim N_d(\mathbf{0}, \Sigma_d(\rho))$ as in Remark 2.1, and consider the vector of $\rho$-values $\phi = (0, 0.4, 0.8, 0.95, 0.99, 1)$. Then compute $I \equiv I(\phi) = (I(\phi_1), \ldots, I(\phi_6))$.

2: By definition, $I(0) = 0$ and $I(1) = 1$.

3: Define $I^* = \varphi_d(I)$.

Based on the values of $I^*$ we define four levels of dependence.

**Definition 4.1.** Let $X \in \mathcal{X}^d$. We say that $X$ has a weak, mild, strong or very strong dependence structure if $I^*$ belongs to $[0, 0.25)$, $[0.25, 0.5)$, $[0.5, 0.75)$ or $[0.75, 1]$ respectively.

**Remark 4.1.** Via numerical study, we observe that when the underlying distribution is $N_2(\mathbf{0}, \Sigma_2(\rho))$ (see Vexler et al., 2019) or $N_3(\mathbf{0}, \Sigma_3(\rho))$ (see Table III.1, Section III of the Supplementary Material) the index $I$ is an increasing function of $\rho \in [0, 1]$. We conjecture that this is true whenever the underlying distribution is $N_d(\mathbf{0}, \Sigma_d(\rho))$ for any value of $d$.

5. Investigation of the index in view of the dependence axioms

Here we investigate the classes of random vectors in which Axioms $A_1$ to $A_8$ and $A'_1$ to $A'_4$ hold; notice that Axiom $A_7$ is only mentioned in the case $d = 2$.

For Axioms $A_3$ and $A'_3$ we define the class of random vectors $\mathcal{X}_1^d$, while for Axioms $A_5$ and $A'_5$ we define the class $\mathcal{X}_2^d$. We first state the conditions that determine these classes of random vectors.

$C_1$: One can detect at least a subscript $j = 1, \ldots, 2^d$ for which the Kendall function of $X_j \in \mathcal{R}(X)$ is such that $K_j(t) - K_H(t) \geq 0$ or $\leq 0$ for all $0 \leq t \leq 1$. That is, we can find at least one Kendall plot that does not cross the diagonal line.

$C_2$: One can detect at least two subscripts $j = 1, \ldots, 2^d$ for which the functions of $X_j \in \mathcal{R}(X)$ are such that $K_j(t) - K_H(t) \leq 0$ or $\geq 0$ for all $0 \leq t \leq 1$. That is, we can find at least two Kendall plots that lie below the diagonal line.

Consider now the following subclasses of $\mathcal{X}_0^d$:

$$\mathcal{X}_1^d \equiv \{ X \in \mathcal{X}_0^d : C_1 \text{ holds} \}; \quad \mathcal{X}_2^d \equiv \{ X \in \mathcal{X}_0^d : C_2 \text{ holds} \}.$$

Note that $\mathcal{X}_2^d \subset \mathcal{X}_1^d \subset \mathcal{X}_0^d$. Moreover, $\mathcal{X}_2^d$ is not an empty set; for each continuous rv $X$, the random vector $(X, \ldots, X)^T$ on $\mathbb{R}^d$ belongs to $\mathcal{X}_2^d$ because its Kendall function is $K(t) = t$, $0 \leq t \leq 1$, see the proof of Equation (4.2) in Appendix D, which lies below of $K_H(t)$, $0 \leq t \leq 1$ (set $g(t) = K_H(t) - t$, $0 \leq t \leq 1$, and observe that $g$ is a concave function with $g(0) = g(1) = 0$). In Section 6 we present Algorithm 6.1, that one can use to check whether the data come from a distribution that belongs in $\mathcal{X}_1^d$ or $\mathcal{X}_2^d$.

We are now in a position to investigate Axioms $A_1$ to $A_8$ as well as $A'_1$ to $A'_4$.

**Proposition 5.1.** For the index of dependence $I$ defined by (4.3) we have: (a) Axioms $A_1$, $A_2$, $A_6$, $A_7$ and $A_8$ hold in $\mathcal{X}_1^d$; (b) Axiom $A_3$ holds in $\mathcal{X}_1^d$; (c) Axioms $A_5$ and $A_8$ hold in $\mathcal{X}_1^d$.

**Proposition 5.2.** For the index of dependence $I$ defined by (4.3) we have: (a) Axiom $A'_1$ holds in $\mathcal{X}_1^d$; (b) Axioms $A'_2$ and $A'_4$ hold in $\mathcal{X}_0^d$; (c) Axioms $A'_3$ holds in $\mathcal{X}_2^d$. 


Remark 5.1. (a) Propositions 5.1 and 5.2 say that $I$ is a measure of dependence for the structure of random vectors in $\mathcal{X}^d_2$, in both Rényi’s as well as Móri and Székely’s definitions.

(b) The standardized index $I^* = \varphi_d(I)$ is also a measure of dependence for the structure of random vectors in $\mathcal{X}^d_2$, in both Rényi’s as well as Móri and Székely’s definitions, whenever $\varphi_d$ is a bijective increasing mapping.

(c) $I$ and $I^*$ are indices of the dependence structure for random vectors in $\mathcal{X}^d_0$.

(d) Consider a random vector $X \sim F$ in $\mathcal{X}^d_2$, and suppose its marginal cdfs are $F_i(t) = \mathbb{P}(X_i \leq t)$, $i = 1, \ldots, d$. Applying the probability integral transform to each $X_i$, the random vector $U = (U_1, \ldots, U_d)^T \equiv \left(F_1(X_1), \ldots, F_d(X_d)\right)^T$ has marginals that are uniformly distributed on $[0, 1]$. The copula of $X$ is defined as the cdf of $U$, $C(u) = \mathbb{P}(U \leq u)$; notice that the copula $C$ contains all information on the dependence structure between the components of $X$. Since Axiom $A_6$ holds, see Proposition 5.1a, we have that the index $I$, and hence $I^*$, is a copula-based index of dependence; namely, $I_F = I_C$ and $I_F^* = I_C^*$.

6. Estimating AUK, $I$, $I^*$

Suppose that a random sample $X_i, i = 1, \ldots, n$, has been drawn from a $d$-dimensional distribution $F$. For each $s_j \in S^d, j = 1, \ldots, d$, set $X_i = \text{diag}(s_j)X_i, i = 1, \ldots, n$, which is a random sample from $X_j = \text{diag}(s_j)X$; and let $F_{nj}$ denote the empirical cumulative distribution function (ecdf) of $X_{ij} = \text{diag}(s_j)X_i, i = 1, \ldots, n$,

$$F_{nj}(x) = \frac{1}{n} \sum_{i=1}^{n} 1_{\{X_{ij} \leq x\}}, \quad x \in \mathbb{R}^d;$$

consider the corresponding empirical Kendall distribution function

$$\mathbb{K}_{nj}(t) = \frac{1}{n} \sum_{i=1}^{n} 1_{\{F_{nj}(X_{ij}) \leq t\}}, \quad t \in [0, 1]. \quad (6.1)$$

Let $\hat{\mathbb{K}}_{nj} \sim \mathbb{K}_{nj}, j = 1, \ldots, d$. By virtue of the formula $\text{AUK} = \mathbb{E}\{1 - K_{II}(T)\}$, we can estimate the AUKs in a nonparametric manner via the statistic

$$\hat{\text{AUK}}_{nj} = \mathbb{E}\{1 - K_{II}(\hat{T}_{nj})\} = 1 - \frac{1}{n} \sum_{i=1}^{n} K_{II}(\hat{T}_{nij}),$$

where $\hat{T}_{nij} = F_{nj}(X_{ij}), i = 1, \ldots, n$. Then, we get nonparametric estimations of $I$ and $I^*$, $\hat{I}_n$ and $\hat{I}^*_n$, respectively, by

$$\hat{I}_n \doteq \left(2^{d-2} - 2^{1-d} + 2^{1-2d}\right)^{-1/2} \|\hat{D}_n - A\| \quad \text{and} \quad \hat{I}^*_n \doteq \varphi_d(\hat{I}_n),$$

where $\hat{D}_n = (\hat{\text{AUK}}_{n1}, \ldots, \hat{\text{AUK}}_{nd})^T$.

We are interested in studying the consistency of the proposed estimators. Theorem C.3 presents the consistency of $\mathbb{K}_{nj}$s, see in Appendix. Based on this result, in Appendix D, we prove that $\hat{\text{AUK}}_{nj}$ is a strongly consistent estimator of $\text{AUK}_j$, that is,

$$\hat{\text{AUK}}_{nj} \xrightarrow{as} \text{AUK}_j \quad \text{for all} \ j = 1, \ldots, 2^d; \quad (6.2)$$
hence, using the continuous mapping theorem, \( \hat{I}_n \xrightarrow{d} I \) as well as \( \hat{I}_n^* \xrightarrow{d} I^* \), that is, both \( \hat{I}_n \) and \( \hat{I}_n^* \) are strongly consistent estimators of \( I \) and \( I^* \) respectively.

In addition to the consistency of \( \hat{I}_n \) and \( \hat{I}_n^* \), Theorem C.3 gives a graphical method for checking if the underlying data distribution belongs in \( \chi_1^d \) or \( \chi_2^d \). Algorithm 6.1 enables one to carry out such a check.

**Algorithm 6.1** Algorithm for checking if the underlying data distribution belongs in \( \chi_1^d \) or \( \chi_2^d \).

1. Compute \( K_H \) and \( K_{nj} \), \( j = 1, \ldots, 2^d \), using the formulas (4.1) and (6.1).
2. Visualize the curves \((K_{nj}(t), K_H(t)), 0 \leq t \leq 1, \) for all \( j = 1, \ldots, 2^d \).
3. **Decision Rule:**
   - If there is at least one curve that does not cross the diagonal line, then the data set is from \( \chi_1^d \).
   - If there are at least two curves that lie below the diagonal line, then the data set is from \( \chi_2^d \).

7. **Testing joint or mutual dependence**

In this section, we investigate the null hypothesis of total independence. Let \( \hat{AUK}_n \) be the estimator of AUK of the original/non-rotated data distribution; namely, the \( \hat{AUK}_{nj} \) that corresponds to \( \mathbf{s}_j = (+, \ldots, +)^T \in S^d \).

We are interested in testing the hypothesis

\[
H_0: F(x) = F_1(x_1) \cdots F_d(x_d) \quad \text{versus} \quad H_1: F(x) \neq F_1(x_1) \cdots F_d(x_d).
\]

For testing this hypothesis, we propose the statistic

\[
z_n = \sqrt{n}(\hat{AUK}_n - 1/2)/\sigma_H,
\]

where the true value of the standard deviation \( \sigma_H \) can be computed explicitly through the relations (7.1b), (7.2a)–(7.2c) given below. In the case \( d = 2 \), we have computed \( \sigma_H = (19/432)^{1/2} \) — see Appendix B.

We now prove the asymptotic normality of the \( \hat{AUK}_n \) under this null hypothesis. The result is derived through the asymptotic behavior of a modified Kendall process that originally was defined by Genest and Rivest (1993) as well as the results given by Barbe et al. (1996). For easy reference, we state these results in Appendix B. In Appendix B, we prove that

\[
\sqrt{n}(\hat{AUK}_n - 1/2) \xrightarrow{d} \mathcal{N}(0, \sigma_H^2), \quad \text{with} \quad \sigma_H^2 = \int_0^1 \int_0^1 k_H(s,t)I_H(s,t) \, ds \, dt = 2 \int_0^1 \int_0^1 k_H(s)I_H(s,t)k_H(t) \, ds \, dt, \quad (7.1b)
\]

where

\[
I_H(s,t) = K_H(s \land t) - K_H(s)K_H(t) + k_H(s)k_H(t)R_H(s,t) - k_H(t)Q_H(s,t) - k_H(s)Q_H(t,s),
\]

\[
Q_H(s,t) = (s \land t) \sum_{i=1}^d \frac{\ln(1/(s \land t))}{i!} - tK_H(s)\quad (7.2a)
\]

where \( u \land v \) denotes the componentwise minimum between \( u \) and \( v \),

\[
Q_H(s,t) = (s \land t) \sum_{i=1}^d \frac{\ln(1/(s \land t))}{i!} - tK_H(s)\quad (7.2b)
\]
To compute the exact value of the true standard deviation \( \sigma \), since standard deviation \( \sigma \) of the standard bivariate uniform distribution. For a sample of size \( n \), the proposed independence test is

\[
\sum_{i=1}^{d} \min \left\{ \ln(s)U_{1}^{(i)}, \ln(t)U_{2}^{(i)} \right\}) - st, \tag{7.2c}
\]

where \( U_j = \left( U_j^{(1)}, \ldots, U_j^{(d)} \right)^T \), \( j = 1, 2 \), are independent random vectors uniformly distributed on the simplex in \([0, 1]^d\).

Therefore, at significance level \( \alpha \), the asymptotic rejection region is \( R = \{ |z_n| > z_{\alpha/2} \} \), where \( z_\alpha \) is the upper \( \alpha \)-point of the standard normal distribution. Hereafter, we call this test as AUK test for total independence or simply AUK independence test.

We investigate further the proposed independence test in the bivariate case. The distribution of \( z_n \) is asymptotically \( N(0, 1) \). Table 7.1 contains the 90th, 95th and 99th percentiles of the empirical distribution of \( |z_n| \) generated by simulation of \( 10^5 \) samples of various sizes \( n \) from the standard bivariate uniform distribution. For a sample of size \( n \), the proposed \( \alpha \)-level independence test is

\[
\text{reject the total independence if } |z_n| > p_{d=2,n;1-\alpha},
\]

where \( p_{d=2,n;1-\alpha} \) is given in Table 7.1 noting that \( p_{d=2,n;1-\alpha} = z_{\alpha/2} \).

**Table 7.1**: Empirical percentiles of the distribution of \(|z_n|\) for random samples of size \( n \) drawn standard bivariate uniform distribution.

| \( n \) | 30 | 50 | 70 | 100 | 150 | 200 | 300 | 400 | 500 | 750 | 1000 | \( \infty \) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| \( p_{d=2,n;0.90} \) | 2.30 | 2.11 | 2.01 | 1.93 | 1.84 | 1.79 | 1.74 | 1.72 | 1.71 | 1.68 | 1.67 | 1.65 |
| \( p_{d=2,n;0.95} \) | 2.62 | 2.44 | 2.34 | 2.25 | 2.17 | 2.12 | 2.06 | 2.05 | 2.03 | 2.01 | 1.98 | 1.96 |
| \( p_{d=2,n;0.99} \) | 3.19 | 3.05 | 2.95 | 2.87 | 2.78 | 2.75 | 2.68 | 2.67 | 2.65 | 2.63 | 2.60 | 2.57 |

Suppose we have a \( d \)-dimensional continuous cdf \( G \) such that the corresponding AUK \( \neq 1/2 \). Due to consistency, see (6.2), and using the continuous mapping theorem, we have that \(|AUK_{\alpha} - 1/2| \xrightarrow{as} |AUK - 1/2| \neq 0 \); therefore, the statistic \(|z_n|\) diverges almost surely to infinity; thus, if the alternative hypothesis is \( F = G \), the power of the test goes to 1. Consequently, the proposed AUK-based test for total independence is consistent against the set of alternatives of \( d \)-dimensional continuous cdfs with AUK \( \neq 1/2 \).

We now offer an alternative way to compute the standard deviation \( \sigma_H \). Recall that the true standard deviation \( \sigma_H \) can be computed exactly through the relations (7.1b) and (7.2a)–(7.2c). Since \( U_i \)'s are uniformly distributed on the \( d \)-dimensional probability simplex, it is rather laborious to compute the exact value of the true standard deviation \( \sigma_H \) for general values of \( d \); the difficulty is due to the calculations on the simplex. Thus, we provide a Monte Carlo based algorithm for estimating the standard deviation \( \sigma_H \).

Using Algorithm 7.1 with \( r = 10000 \) and \( n = 50000 \), we offer estimations \( \hat{\sigma}_H \) for the cases \( d = 2, \ldots, 10 \). Observe that in the case \( d = 2 \) the estimated value \( \hat{\sigma}_H = 0.20988 \) is very closed to the true value \( \sigma_H = (19/432)^{1/2} \approx 0.20972 \) \((\hat{\sigma}_H / \sigma_H = 1.000779)\).

To make our test practical for small sample sizes \( n \), we present an algorithm to approximate the percentiles of the distribution of \(|z_n|\). Suppose \( X_1, \ldots, X_n \) is a random sample of size \( n \) from a \( d \)-dimensional continuous distribution \( F \) and we are interested in testing the null hypothesis of total independence. We use the statistic \( z_n \); and the rejection region of the null hypothesis at \( \alpha \)-significant level is \(|z_n| > p_{d,n;1-\alpha}\), where \( p_{d,n;1-\alpha} \) denote the \((1 - \alpha)\)-percentile point.
Algorithm 7.1 Approximation of the standard deviation $\sigma_H$.

1: Let $d$ be the dimension of interest. Consider a $d$-dimensional continuous distribution with independent components; without loss of generality, consider the standard $d$-dimensional uniform distribution, say $U$, that is the distribution with density $f_U(u) = u_1 \cdots u_d, u = (u_1, \ldots, u_d)^T \in [0, 1]^d$.
2: Let $n$ and $r$ be two large positive integer numbers. Generate $r$ independent random samples of size $n$ from $F_U$, say $U_{i,j}, i = 1, \ldots, n$ and $j = 1, \ldots, r$.
3: Based on the $j$th random sample, $U_{i,j}, i = 1, \ldots, n$, compute the AUK$_j$ and then $A_i = \sqrt{n}(\text{AUK}_j - 1/2), j = 1, \ldots, r$.
4: Define $\hat{\sigma}_H$ to be the sample standard deviation of $A_1, \ldots, A_r$.

Table 7.2: Monte Carlo estimate of $\sigma_H$, using Algorithm 7.1 with $r = 10000$ and $n = 50000$, when $d = 2, \ldots, 10$.

| $d$ | 2   | 3   | 4   | 5   | 6   | 7   | 8   | 9   | 10  |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| $\hat{\sigma}_H$ | 0.20988 | 0.19383 | 0.16254 | 0.12511 | 0.09407 | 0.06853 | 0.04912 | 0.03395 | 0.02377 |

of the distribution of $|z_n|$. If the sample size $n$ is large, that is $n > \max\{1000, 100d\}$, then $p_{d,n;1-\alpha} = p_{d,\infty;1-\alpha} = z_{n/2}$ else we approximate this percentile point using Algorithm 7.2.

Algorithm 7.2 Approximation of the percentiles of the distribution of $|z_n|$ for small sample sizes.

1: Consider the standard $d$-dimensional uniform distribution, say $U$, that is the distribution with density $f_U(u) = u_1 \cdots u_d, u = (u_1, \ldots, u_d)^T \in [0, 1]^d$.
2: For a large positive integer $r$, generate $r$ independent random samples of size $n$ from $F_U$, say $U_{i,j}, i = 1, \ldots, n$ and $j = 1, \ldots, r$.
3: Based on the $j$th random sample, $U_{i,j}, i = 1, \ldots, n$, compute the corresponding $|z_n|$-value, say $|z_n|$, $j = 1, \ldots, r$.
4: Define $p_{d,n;1-\alpha}$ to be the $(1 - \alpha)$-percentile point of the distribution of the values $|z_n|, \ldots, |z_n|$.

Notice that if the dimension $d$ is small, for example $d \leq 10$, and since $n \leq 1000$, the number of Monte Carlo repetitions $r = 10^4$ is sufficient for an accurate approximation of $p_{d,n;1-\alpha}$ using Algorithm 7.2 with low computational cost.

Remark 7.1. For the estimation of AUK we used the relationship AUK = $E(1 - K_H(T))$. The study of the asymptotic behavior of this estimator, i.e. the estimator AUK$_n$ that is presented in Section 7, as well as its limit distribution under the hypothesis of total independence go through the Kendall process and the functional delta-method. Another way for estimating the AUK is via use of the formula AUK = $P(T \leq \Pi)$; then, similar to Pfister et al. (2018), a kernel-based estimation of AUK seems to be feasible and then the use of U- or V-statistics theory allows one to obtain the asymptotic distribution of the statistic.

We now discuss the possibility of a general AUK-based hypothesis test, i.e. $H_0$: AUK = $a_0$ versus $H_1$: AUK $\neq a_0$ for a fixed constant $a_0 \in [2^{-d}, 1]$; until now we have only studied the case $a_0 = 1/2$. We first present the following example.

Example 7.1. Let $X = (X_1, X_2)^T$ be a bivariate rv uniformly distributed on the circumference of the circle centered at the origin with a radius of 1. The ordinary Kendall cdf of $X$ is $K_X(t) = 1_{[0,\sqrt{1/2}]}(t + 1/4) + 1_{[1/2,\sqrt{1}]})$ (see Vexler et al., 2019, p. 426) and so AUK$_X = 11/16 - \ln(2)/4$. Consider now the bivariate normal rv $Z = (Z_1, Z_2)^T \sim \mathcal{N}_2\left(0, \begin{pmatrix} 1 & -0.073622 \\ -0.073622 & 1 \end{pmatrix} \right)$. A numerical
computation gives $\text{AUK}_Z = \text{AUK}_X$. One can verify that hypotheses $H_1$ and $H_2$, see Appendix B, hold and so Theorem B.1 also holds for the random vector $Z$. On the other hand, hypothesis $H_1$ does not hold for the random vector $X$; hence, the statement of Theorem B.1 is not true for this case.

In view of the above example, we conclude that it is not possible to construct general AUK-based hypotheses tests, as these are described at the beginning of the paragraph.

8. Simulation study

In this section, we present a simulation study designed with two specific goals in mind. First, we are interested in understanding the behavior of the proposed indices of dependence as a function of the sample size, the type and the degree of dependence that a data set exhibits.

Our second goal relates to the evaluation of the performance of the proposed tests of independence, in terms of level of significance and power of the tests. In what follows, we present evidence of this performance by simulating data from a range of distributions and hence, dependence structures.

8.1. The estimators $\hat{I}_n$ and $\hat{I}^*_n$

We first present a simulation study for the proposed empirical dependence indices.

8.1.1. Normal case

We simulate $r = 1000$ random samples of size $n = 100, 200, 500, 1000$ from the trivariate normal $\mathcal{N}_3(\mathbf{0}, \Sigma_3(\rho))$ distribution (for the variance/covariance matrix $\Sigma_3(\rho)$ see (2.2a)), for various values of $n$ and $\rho = (\rho_{12}, \rho_{13}, \rho_{23})$. Based on these samples we compute the averages (and the Monte Carlo mean square error) of $\hat{I}_n$ and $\hat{I}^*_n$; Table 8.1 presents the results. In all cases the empirical mean square error of the indices tends to zero as $n$ goes to infinity, confirming the consistency of the estimation for both $\hat{I}_n$ and $\hat{I}^*_n$ estimators. Figure 8.1 presents illustrations of graphs of a single sample $\{X_1, \ldots, X_n\}$ of size $n = 1000$ drawn from these distributions. The standardized index $I^*$ indicates that, see Definition 4.1, the normal distributions with $(\rho_{12}, \rho_{13}, \rho_{23}) \in \{(0,0,0), (-0.1,-0.1,0.2)\}$ have a weak dependence level, that the ones with $(\rho_{12}, \rho_{13}, \rho_{23}) \in \{(-0.3,-0.3,-0.3), (0.2,0.3,0.4)\}$ have a mild dependence level, while the distributions with $(\rho_{12}, \rho_{13}, \rho_{23}) \in \{(0.7,0.5,0), (0.2,-0.8,0)\}$ have a strong dependence level, and that normal distributions with $(\rho_{12}, \rho_{13}, \rho_{23}) \in \{(1,1,1), (-0.5,-0.5,-0.5), (0.1,0.2,-0.9)\}$ have a very strong dependence level. On the other hand, the case $(\rho_{12}, \rho_{13}, \rho_{23}) = (-0.5,-0.5,0.5)$ lies in the frontier between mild and strong dependence level.

8.1.2. Non-normal cases

We simulate $r = 1000$ random samples of size $n$ from some trivariate Archimedean copulas (Clayton, Frank, Gumbel and Joe) as well the non-Archimedean copula Farlie-Gumbel-Morgenstern (F-G-M). For the parameters of these Archimedean copulas see the R-package ‘copula’ (see, also, Brechmann and Schepsmeier, 2013); regarding data generation from the F-G-M copula see Section I of the Supplementary Material. Based on these samples we compute the averages (and
Table 8.1: Monte Carlo averages (and associated mean square error) of $I$ and $I^*$; the number of Monte Carlo repetitions is $r = 1000$. Random samples of size $n$ are drawn from different $\mathcal{N}_3(0, \Sigma_3(\rho))$ distributions, for various values of $n$ and $\rho = (\rho_{12}, \rho_{13}, \rho_{23})$.

| $n$  | 100      | 200      | 500      | 1000     |
|------|----------|----------|----------|----------|
| $I_n$ | 0.97(0.009) | 0.59(0.004) | 0.32(0.001) | 0.22(0.000) |
| $I'_n$ | 0.186(0.035) | 0.109(0.012) | 0.057(0.003) | 0.037(0.001) |

(a) $(\rho_{12}, \rho_{13}, \rho_{23}) = (0, 0, 0)$; true values of indices: $I = 0$ and $I^* = 0.$

(b) $(\rho_{12}, \rho_{13}, \rho_{23}) = (1, 1, 1)$; true values of indices: $I = 1$ and $I^* = 1.$

| $n$  | 100      | 200      | 500      | 1000     |
|------|----------|----------|----------|----------|
| $I_n$ | 0.739(0.068) | 0.833(0.028) | 0.912(0.008) | 0.947(0.003) |
| $I'_n$ | 0.985(0.000) | 0.994(0.000) | 0.998(0.000) | 0.999(0.000) |

(c) $(\rho_{12}, \rho_{13}, \rho_{23}) = (-0.5, -0.5, 0.5)$; true values of indices: $I = 0.243$ and $I^* = 0.5.$

(d) $(\rho_{12}, \rho_{13}, \rho_{23}) = (-0.5, -0.5, -0.5)$; true values of indices: $I = 0.554$ and $I^* = 0.92.$

| $n$  | 100      | 200      | 500      | 1000     |
|------|----------|----------|----------|----------|
| $I_n$ | 0.222(0.001) | 0.229(0.000) | 0.236(0.000) | 0.239(0.000) |
| $I'_n$ | 0.459(0.004) | 0.472(0.002) | 0.487(0.001) | 0.493(0.000) |

(e) $(\rho_{12}, \rho_{13}, \rho_{23}) = (0.10, 0.20, -0.90)$; true values of indices: $I = 0.461$ and $I^* = 0.842.$

(f) $(\rho_{12}, \rho_{13}, \rho_{23}) = (0.70, 0.50, 0)$; true values of indices: $I = 0.363$ and $I^* = 0.717.$

| $n$  | 100      | 200      | 500      | 1000     |
|------|----------|----------|----------|----------|
| $I_n$ | 0.382(0.006) | 0.413(0.002) | 0.438(0.000) | 0.449(0.000) |
| $I'_n$ | 0.745(0.010) | 0.787(0.003) | 0.818(0.000) | 0.829(0.000) |

(g) $(\rho_{12}, \rho_{13}, \rho_{23}) = (0.20, -0.80, 0)$; true values of indices: $I = 0.334$ and $I^* = 0.671.$

(h) $(\rho_{12}, \rho_{13}, \rho_{23}) = (-0.30, -0.30, -0.30)$; true values of indices: $I = 0.204$ and $I^* = 0.420.$

| $n$  | 100      | 200      | 500      | 1000     |
|------|----------|----------|----------|----------|
| $I_n$ | 0.291(0.002) | 0.307(0.001) | 0.322(0.000) | 0.327(0.000) |
| $I'_n$ | 0.594(0.007) | 0.624(0.003) | 0.651(0.001) | 0.659(0.000) |

(i) $(\rho_{12}, \rho_{13}, \rho_{23}) = (0.20, 0.30, 0.40)$; true values of indices: $I = 0.159$ and $I^* = 0.324.$

(j) $(\rho_{12}, \rho_{13}, \rho_{23}) = (-0.10, -0.10, 0.20)$; true values of indices: $I = 0.074$ and $I^* = 0.138.$

| $n$  | 100      | 200      | 500      | 1000     |
|------|----------|----------|----------|----------|
| $I_n$ | 0.165(0.001) | 0.156(0.000) | 0.156(0.000) | 0.157(0.000) |
| $I'_n$ | 0.336(0.003) | 0.316(0.002) | 0.317(0.001) | 0.319(0.000) |

| $n$  | 100      | 200      | 500      | 1000     |
|------|----------|----------|----------|----------|
| $I_n$ | 0.114(0.002) | 0.089(0.000) | 0.077(0.000) | 0.075(0.000) |
| $I'_n$ | 0.224(0.009) | 0.170(0.002) | 0.146(0.001) | 0.141(0.000) |
the Monte Carlo mean square error) of $\hat{I}$ and $\hat{I}^*$; Tables 8.2 and 8.3 present the results for the Archimedean and F-G-M copula cases respectively; note that the copulas $\tilde{C}_\theta$ or $\tilde{C}_{-\theta}$, $\theta \in [0,1]$, see (L.1a), produce the same value of the index $I$, and so the index $I^*$ [the same is true for the copulas $\tilde{C}_\theta$ or $\tilde{C}_{-\theta}$, $\theta \in [0,1]$, see (L.1a)]. For all cases the results again empirically verify the consistency of the estimation for both $\hat{I}$ and $\hat{I}^*$ estimators. Figure 8.1 presents illustrations of graphs of a single sample $\{X_1, \ldots, X_n\}$ of size $n = 1000$ drawn from these distributions. The standardized index $I^*$ indicates that the Archimedean copulas Clayton[2], Frank[4], Gumbel[2] and Joe[2] have a strong dependence level, while those corresponding to Clayton[5], Frank[8], Gumbel[4] and Joe[5] have a very strong dependence level.

**General Observations:** 1) As the sample size increases the bias decreases; 2) The highest percentage of bias, across all sample sizes, is observed in the case of complete dependence $(\rho_{12}, \rho_{13}, \rho_{23}) = (1, 1, 1)$. In this case $I = I^* = 1$. The observed bias (defined as $\hat{I} - I$) goes from $-26.1\%$ ($n = 100$, $I$) to $-0.1\%$ ($n = 1000$, $I^*$).

### 8.2. Evaluating the proposed test of independence

We now investigate the performance of the test of independence, introduced in Section 7. Performance will be investigated in terms of both, level and power. The dimensions of the generated data are $d = 2, 6$. We simulate normal and non-normal data distributions. We choose to work with the bivariate case because it allows comparisons between numerical results produced and the shape of the data distributions. Figure 8.4 illustrates graphically the shape of a single random sample of size 1000 drawn from bivariate non-normal distributions we work with, with various degrees of dependence.

The results shown in Table 8.4 verify the consistency of the proposed independence test, in the bivariate normal case. As the sample size increases to infinity the power of the test increases to 1. The rate of convergence of the power to 1 depends on the strength of the association between the components of the random vector.

Figure 8.1: Scatterplots: Random samples of size $n = 1000$ are drawn from the the trivariate normal distributions $N_3(0, \Sigma_3(\rho))$ for various correlation values of $\rho = (\rho_{12}, \rho_{13}, \rho_{23})$.
AUK-based index for measuring and testing the joint dependence of a random vector

Table 8.2: Monte Carlo averages (and associated mean square error) of $\mathcal{I}$ and $\mathcal{I}^*$; the number of Monte Carlo repetitions is $r = 1000$. Random samples of size $n$ are drawn from different Clayton, Frank, Gumbel and Joe Archimedean copulas, for various values of $n$. For each case, the parameter is shown in curly brackets, and the true values of $\mathcal{I}$ and $\mathcal{I}^*$ are denoted in the caption.

(a) Clayton (2); $\mathcal{I} = 0.362$, $\mathcal{I}^* = 0.715.$

| $n$ | \( \hat{L}_n \) | \( \hat{L}_n^* \) |
|-----|-----------------|-----------------|
| 100 | 0.315(0.003)    | 0.637(0.008)    |
| 200 | 0.332(0.001)    | 0.666(0.003)    |
| 500 | 0.348(0.000)    | 0.694(0.001)    |
| 1000| 0.354(0.000)    | 0.703(0.000)    |

(b) Clayton (5); $\mathcal{I} = 0.528$, $\mathcal{I}^* = 0.903.$

| $n$ | \( \hat{L}_n \) | \( \hat{L}_n^* \) |
|-----|-----------------|-----------------|
| 100 | 0.442(0.008)    | 0.821(0.007)    |
| 200 | 0.479(0.003)    | 0.860(0.002)    |
| 500 | 0.507(0.001)    | 0.886(0.000)    |
| 1000| 0.517(0.000)    | 0.893(0.000)    |

(c) Frank (4); $\mathcal{I} = 0.263$, $\mathcal{I}^* = 0.543.$

| $n$ | \( \hat{L}_n \) | \( \hat{L}_n^* \) |
|-----|-----------------|-----------------|
| 100 | 0.241(0.001)    | 0.497(0.000)    |
| 200 | 0.247(0.001)    | 0.510(0.002)    |
| 500 | 0.255(0.000)    | 0.527(0.001)    |
| 1000| 0.260(0.000)    | 0.535(0.000)    |

(d) Frank (8); $\mathcal{I} = 0.410$, $\mathcal{I}^* = 0.783.$

| $n$ | \( \hat{L}_n \) | \( \hat{L}_n^* \) |
|-----|-----------------|-----------------|
| 100 | 0.354(0.004)    | 0.702(0.007)    |
| 200 | 0.378(0.001)    | 0.739(0.002)    |
| 500 | 0.396(0.000)    | 0.764(0.001)    |
| 1000| 0.403(0.000)    | 0.774(0.000)    |

(e) Gumbel (2); $\mathcal{I} = 0.361$, $\mathcal{I}^* = 0.714.$

| $n$ | \( \hat{L}_n \) | \( \hat{L}_n^* \) |
|-----|-----------------|-----------------|
| 100 | 0.316(0.003)    | 0.638(0.007)    |
| 200 | 0.333(0.001)    | 0.668(0.003)    |
| 500 | 0.348(0.000)    | 0.693(0.001)    |
| 1000| 0.354(0.000)    | 0.703(0.000)    |

(f) Gumbel (4); $\mathcal{I} = 0.562$, $\mathcal{I}^* = 0.925.$

| $n$ | \( \hat{L}_n \) | \( \hat{L}_n^* \) |
|-----|-----------------|-----------------|
| 100 | 0.471(0.008)    | 0.853(0.006)    |
| 200 | 0.509(0.003)    | 0.887(0.002)    |
| 500 | 0.539(0.001)    | 0.910(0.000)    |
| 1000| 0.550(0.000)    | 0.918(0.000)    |

(g) Joe (2); $\mathcal{I} = 0.267$, $\mathcal{I}^* = 0.549.$

| $n$ | \( \hat{L}_n \) | \( \hat{L}_n^* \) |
|-----|-----------------|-----------------|
| 100 | 0.240(0.001)    | 0.495(0.005)    |
| 200 | 0.248(0.001)    | 0.512(0.003)    |
| 500 | 0.258(0.000)    | 0.532(0.001)    |
| 1000| 0.262(0.000)    | 0.540(0.000)    |

(h) Joe (5); $\mathcal{I} = 0.496$, $\mathcal{I}^* = 0.876.$

| $n$ | \( \hat{L}_n \) | \( \hat{L}_n^* \) |
|-----|-----------------|-----------------|
| 100 | 0.416(0.007)    | 0.790(0.008)    |
| 200 | 0.449(0.002)    | 0.829(0.003)    |
| 500 | 0.475(0.001)    | 0.857(0.001)    |
| 1000| 0.484(0.000)    | 0.866(0.000)    |
Figure 8.2: Scatterplots: Random samples of size $n = 1000$ are drawn from Clayton, Frank, Gumbel and Joe trivariate copulas for various values of their parameters; the parameter value is shown in curly brackets.

Figure 8.3: Scatterplots: Random samples of size $n = 1000$ are drawn from different trivariate F-G-M copulas $C_\theta$ and $\tilde{C}_\theta$ for various values $\theta$, see (1.1).
Table 8.3: Monte Carlo averages (and associated mean square error) of $I$ and $I^*$; the number of Monte Carlo repetitions is $r = 1000$. Random samples of size $n$ are drawn from different F-G-M copulas given by (I.1a) or (I.1b), for various values of $\theta$ and $n$. For each case, the true values of $I$ and $I^*$ are denoted in the caption.

(a) $C_{0.5}; I = 0.052, I^* = 0.093.$

| $n$   | 100  | 200  | 500  | 1000 |
|-------|------|------|------|------|
| $\hat{I}_n$ | .105(.003) | .075(.001) | .059(.000) | .055(.000) |
| $\hat{I}_n^*$ | .205(.013) | .141(.003) | .108(.001) | .100(.000) |

(b) $C_{0.7}; I = 0.073, I^* = 0.136.$

| $n$   | 100  | 200  | 500  | 1000 |
|-------|------|------|------|------|
| $\hat{I}_n$ | .110(.002) | .087(.001) | .076(.000) | .074(.000) |
| $\hat{I}_n^*$ | .220(.009) | .167(.002) | .144(.000) | .138(.000) |

(c) $C_{0.9}; I = 0.094, I^* = 0.180.$

| $n$   | 100  | 200  | 500  | 1000 |
|-------|------|------|------|------|
| $\hat{I}_n$ | .123(.001) | .102(.000) | .095(.000) | .094(.000) |
| $\hat{I}_n^*$ | .245(.006) | .198(.002) | .184(.000) | .181(.000) |

(d) $C_{1}; I = 0.104, I^* = 0.203.$

| $n$   | 100  | 200  | 500  | 1000 |
|-------|------|------|------|------|
| $\hat{I}_n$ | .129(.001) | .110(.000) | .105(.000) | .104(.000) |
| $\hat{I}_n^*$ | .257(.004) | .216(.001) | .204(.001) | .203(.000) |

(e) $\tilde{C}_{0.5}; I = 0.020, I^* = 0.033.$

| $n$   | 100  | 200  | 500  | 1000 |
|-------|------|------|------|------|
| $\hat{I}_n$ | .098(.005) | .062(.001) | .037(.000) | .028(.000) |
| $\hat{I}_n^*$ | .189(.021) | .114(.005) | .065(.001) | .049(.000) |

(f) $\tilde{C}_{0.7}; I = 0.028, I^* = 0.048.$

| $n$   | 100  | 200  | 500  | 1000 |
|-------|------|------|------|------|
| $\hat{I}_n$ | .098(.004) | .063(.001) | .041(.000) | .034(.000) |
| $\hat{I}_n^*$ | .190(.016) | .117(.003) | .074(.001) | .059(.000) |

(g) $\tilde{C}_{0.9}; I = 0.035, I^* = 0.062.$

| $n$   | 100  | 200  | 500  | 1000 |
|-------|------|------|------|------|
| $\hat{I}_n$ | .099(.003) | .066(.001) | .046(.000) | .040(.000) |
| $\hat{I}_n^*$ | .192(.012) | .121(.002) | .082(.001) | .071(.000) |

(h) $\tilde{C}_{1}; I = 0.039, I^* = 0.070.$

| $n$   | 100  | 200  | 500  | 1000 |
|-------|------|------|------|------|
| $\hat{I}_n$ | .100(.002) | .067(.001) | .049(.000) | .043(.000) |
| $\hat{I}_n^*$ | .194(.010) | .125(.002) | .088(.001) | .077(.000) |
Table 8.4 also presents a comparison between the proposed test and the test for independence based on the estimated dHSIC statistic. We simulated this test using the R package ‘dHSIC’ by Pfister and Peters (2019). Our test exhibits higher power than the dHSIC-based test for all smaller samples and smaller values of \( \rho \), while it exhibits the same power with the dHSIC-based test for larger samples and values of \( \rho \).

Table 8.4: Empirical rejection rates (%) at the \( \alpha = 5\% \) significance level of AUK and dHSIC independence tests, calculated from 1000 random samples of size \( n \) simulated from \( \mathcal{N}_2(\mathbf{0}, \mathbf{\Sigma}(\rho)) \) distribution for various values of \( \rho \). The first two rows correspond to the empirical level of significance of the tests, while the other ones correspond to the empirical power of the tests.

| \( \rho \) | test | \( n \) | 50 | 100 | 200 | 300 | 500 | 750 | 1000 | 1500 | 2000 |
|--------|------|------|----|----|----|----|----|----|------|------|------|
| 0.0    | AUK  |      | 5.2| 4.9| 4.8| 4.9| 5   | 5.1| 4.9 | 5.1  | 5    |
|        | dHSIC|      | 5.5| 5.9| 5.1| 4.7| 4.3 | 4.7| 5.5 | 5.4  | 5.2  |
| 0.1    | AUK  |      | 14.7| 22 | 32.1| 42.9| 62.1| 76.6| 86.5| 95.2 | 98.2 |
|        | dHSIC|      | 6.8 | 8.1 | 12.7| 15  | 24.2| 37.9| 48.4| 70.3 | 80.3 |
| 0.2    | AUK  |      | 35.2| 54.5| 80.4| 93.3| 99  | 99.9| 100 | 100  | 100  |
|        | dHSIC|      | 13.2| 23.3| 40.1| 59.6| 84.4| 95.8| 99.3| 100  | 100  |
| 0.3    | AUK  |      | 59.2| 84.8| 98.9| 100 | 100 | 100 | 100 | 100  | 100  |
|        | dHSIC|      | 27.7| 43.3| 80.2| 94.8| 99.8| 100 | 100 | 100  | 100  |
| 0.4    | AUK  |      | 84.6| 97.6| 100 | 100 | 100 | 100 | 100 | 100  | 100  |
|        | dHSIC|      | 46.3| 79  | 98.1| 99.9| 100 | 100 | 100 | 100  | 100  |
| 0.5    | AUK  |      | 95.6| 99.9| 100 | 100 | 100 | 100 | 100 | 100  | 100  |
|        | dHSIC|      | 71  | 96.6| 99.9| 100 | 100 | 100 | 100 | 100  | 100  |
| 0.6    | AUK  |      | 99.4| 100 | 100 | 100 | 100 | 100 | 100 | 100  | 100  |
|        | dHSIC|      | 92.6| 100 | 100 | 100 | 100 | 100 | 100 | 100  | 100  |

In view of Table 8.5, we see that the consistency of the proposed test of independence is also verified (as the sample size increases to infinity the power of the test tends to 1).

The case \( U\{C(0, 1)\} \) indicates an exact cyclical dependence (see Figure 8.4a). In this case, the Kendall cdfs are \( K_i(t) = \mathbf{1}_{[0 < t < 1/2]}(t+1/4) + \mathbf{1}_{[1/2 < t < 1]} \) and so AUK \( = 11/16 – \ln(2)/4 \approx 0.514, i = 1, \ldots, 4 \) — see Example 7.1; hence, we compute \( \rho = 0.036 \) and \( \rho^* = 0.074 \). This kind of dependence is not recognized by both AUK-indices. On the other hand, AUK is slightly different than 1/2 and so we expect the power of the proposed test will tend to 1 as \( n \) goes to infinity. Indeed, this happens even in this limited case, see the first row of the Table 8.5; additionally, we have that the empirical power for \( n = 2500 \) and 3000 is 0.941 and 0.986 respectively.

In view of Figure 8.4j, one may incorrectly conclude that there is an association between the marginal distributions of the data distribution. Since \( (1/X, 1/Y) \sim \mathcal{N}_2(\mathbf{0}, \mathbf{I}) \), \( X \) and \( Y \) are independent continuous rvs. Thus, the rate of rejection of the null hypothesis corresponds to the empirical size of the tests. For the same values of \( n \) as in Table 8.5, the empirical size of the AUK independence test calculated from 1000 random samples of size \( n \) results in the following values 0.060, 0.053, 0.060, 0.052, 0.039, 0.052, 0.057, 0.054 and 0.046, that is, the test maintains the nominal level well for all sample sizes; notice that the dHSIC independence test also maintains the nominal level well for all sample sizes.

Table 8.5 presents the empirical power of our proposed test, as well as that of the dHSIC-based test for a variety of sample sizes and distributions. We observe that for distributions, such
Figure 8.4: Scatter plots: Random samples of size $n = 1000$ are drawn from different bivariate distributions.
as Morgenstern\{0.5\}, Plackett\{1.25\}, Plackett\{2\} or \(t_5\) with the reported covariance matrix our proposed test outperforms the dHSIC-based test for smaller samples and it has the same power with dHSIC-based test for larger samples. The case where the dHSIC-based test outperforms our test in terms of power is when the data come from \(U\{C(0, 1)\}\) and spiral data cases 3 and 4. Hence, overall the proposed test either performs competitively or outperforms the kernel-based tests in most of the cases met in practice.

Table 8.5: Empirical power (%) of AUK and dHSIC independence tests, with a nominal significance level of \(\alpha = 0.05\), calculated from 1,000 random samples of size \(n\) simulated from various bivariate non-normal distributions.

| distribution         | test  | \(n\)  | 50  | 100 | 200 | 300 | 500 | 750 | 1000 | 1500 | 2000 |
|----------------------|-------|-------|-----|-----|-----|-----|-----|-----|------|------|------|
| \(U\{C(0, 1)\}\)    | AUK   | 0     | 0.2 | 1.1 | 3.1 | 9.3 | 16.0| 36.9| 69   | 86.1 |      |
|                      | dHSIC | 96.7  | 100 | 100 | 100 | 100 | 100 | 100 | 100  | 100  |      |
| \(\exp(2, 3, 1.3)\)  | AUK   | 97.3  | 100 | 100 | 100 | 100 | 100 | 100 | 100  | 100  | 100  |
|                      | dHSIC | 90.5  | 99.9| 100 | 100 | 100 | 100 | 100 | 100  | 100  | 100  |
| Morgenstern\{0.5\}  | AUK   | 27.4  | 42.7| 66.1| 82.1| 95.8| 99  | 99.9| 100  |      |      |
|                      | dHSIC | 16.6  | 28.1| 50.4| 69.7| 87.9| 97.9| 99.6| 100  |      |      |
| Morgenstern\{5\}    | AUK   | 100   | 100 | 100 | 100 | 100 | 100 | 100 | 100  | 100  | 100  |
|                      | dHSIC | 100   | 100 | 100 | 100 | 100 | 100 | 100 | 100  | 100  | 100  |
| Plackett\{1.25\}    | AUK   | 11.4  | 14.1| 22.9| 29.5| 40.3| 53.2| 68.7| 82   | 90.5 |      |
|                      | dHSIC | 6.3   | 10  | 15.6| 19.7| 26.8| 40.4| 48.9| 67.9 | 80.1 |      |
| Plackett\{2\}       | AUK   | 46.8  | 70.4| 92.1| 98.7| 99.9| 100 | 100 | 100  |      |      |
|                      | dHSIC | 29.6  | 57.2| 86.9| 96.4| 99.8| 100 | 100 | 100  |      |      |
| ALI-HAQ’s\{0.9, 0.5\}| AUK | 100  | 100 | 100 | 100 | 100 | 100 | 100 | 100  | 100  | 100  |
|                      | dHSIC | 100   | 100 | 100 | 100 | 100 | 100 | 100 | 100  | 100  | 100  |
| Gumbel\{0.9\}       | AUK   | 22.1  | 35.3| 52.6| 68.1| 85.9| 95.4| 98.7| 100  |      |      |
|                      | dHSIC | 24.9  | 43.7| 76.1| 90.4| 99.2| 100 | 100 | 100  |      |      |
| \(t_5\) with \(\Sigma = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\) | AUK | 93.4  | 99.9| 100 | 100 | 100 | 100 | 100 | 100  |      |      |
|                      | dHSIC | 76.9  | 97.3| 100 | 100 | 100 | 100 | 100 | 100  |      |      |
| \((X, \varepsilon/X^2)\) see Figure 8.4k | AUK | 100  | 100 | 100 | 100 | 100 | 100 | 100 | 100  |      |      |
|                      | dHSIC | 100   | 100 | 100 | 100 | 100 | 100 | 100 | 100  |      |      |
| \((X, X/\varepsilon)\) see Figure 8.4l | AUK | 100  | 100 | 100 | 100 | 100 | 100 | 100 | 100  |      |      |
|                      | dHSIC | 100   | 100 | 100 | 100 | 100 | 100 | 100 | 100  |      |      |
| Spiral data: case 1  | AUK   | 14.5  | 22.9| 35.6| 44.9| 60.5| 76  | 86.7| 95.4 | 98.8 |      |
|                      | dHSIC | 28.6  | 58.5| 93.7| 99.2| 100 | 100 | 100 | 100  | 100  | 100  |
| Spiral data: case 2  | AUK   | 19.6  | 36.8| 57.2| 71.8| 89.8| 97.8| 99.8| 100  |      |      |
|                      | dHSIC | 10.5  | 19  | 41.4| 69.1| 95.3| 99.9| 100 | 100  |      |      |
| Spiral data: case 3  | AUK   | 8.2   | 7.5 | 11.2| 18.5| 23.7| 33.9| 44.5| 58.9 | 70.4 |      |
|                      | dHSIC | 6.8   | 14.5| 35.7| 62.1| 94.2| 100 | 100 | 100  |      |      |
| Spiral data: case 4  | AUK   | 6.4   | 5.7 | 6.7 | 6.4 | 5.5 | 5.7 | 5.6 | 5.5  | 7.4  |      |
|                      | dHSIC | 35.5  | 65.7| 94.5| 99.9| 100 | 100 | 100 | 100  |      |      |

Table 8.6 provides the time, in seconds, needed to compute our AUK-based and the dHSIC-based tests of independence. This table, clearly shows that the AUK-based test is very fast to compute, requiring, for example, for sample sizes 2000 and dimension 2 only 0.048 sec.
Table 8.6: Speed time (in seconds) of the test-procedures AUK and dHSIC for testing the total independence of a two-variable data set of size $n$.

| $n$  | 200 | 300 | 500 | 750 | 1000 | 1250 | 1500 | 1750 | 2000 |
|------|-----|-----|-----|-----|------|------|------|------|------|
| AUK test | 0.000 | 0.000 | 0.015 | 0.016 | 0.017 | 0.031 | 0.032 | 0.047 | 0.048 |
| dHSIC test | 0.609 | 1.500 | 4.633 | 10.017 | 19.814 | 33.800 | 56.782 | 72.949 | 113.030 |

9. A real data example

Alanine and aspartate aminotransferase, ALT and AST respectively, are markers of liver inflammation. Further, alkaline phosphatase (AP), as well as direct bilirubin (DB) are biomarkers of hepatic injury, measuring inflammation from viral, metabolic or autoimmune causes. Elevation of ALT and AST is referred to as hepatocellular liver injury pattern. In contrast, alkaline phosphatase and bilirubin are increased under conditions that lead to damage of the bile ducts that drain the liver and is referred to as cholestatic injury pattern. These parameters can be elevated due to bile duct obstruction by a stone or tumor or in the case of autoimmune diseases, in which body’s immune system can attack the cholangiocytes. There can also be a mixed hepatocellular and cholestatic pattern. In the evaluation of liver disease, the pattern of injury guides the clinician toward the potential etiology and assists in directing the etiologic evaluation. Table IV.1 of the Supplementary Material presents a data set that contains measurements of (DB, AST, ALT, AP) on 208 patients with liver disease.

We estimate the indices $I$ and $I^*$ of the biomarkers’ vector as $\hat{I}_{208}(\text{DB, AST, ALT, AP}) = 0.2546$ and $\hat{I}^*_{208}(\text{DB, AST, ALT, AP}) = \varphi(0.2546) \approx 0.561$, where the approximate polynomial $\varphi_4$ has been find by applying Algorithm 4.1. According to Definition 4.1, we conclude that the random vector of the biomarkers DB, AST, ALT and AP has a strong mutual dependence at the level 56.1%.

We next investigate the dependence structure of the random vector of these biomarkers. To accomplish this goal, we compute the indices $\hat{I}_n$ and $\hat{I}^*_n$ for all 2- and 3-dimensional sub-vectors of the vector of these biomarkers; Tables 9.1a and 9.1b present these results. According to Definition 4.1, we extract the following conclusions. Both (DB, AST, ALT) and (AST, ALT, AP) have strong mutual dependence of level 65.1% and 60.5% respectively, (DB, AST, AP) has a mild mutual dependence of level 29.4% and (DB, ALT, AP) has a weak mutual dependence of level 23.2%; for the bidimensional sub-vectors, we have that (AST, ALT) has a very strong mutual dependence of level 79.2%, (DB, AST) has a mild mutual dependence of level 35.4%, and (DB, ALT), (DB, AP), (AST, AP) and (ALT, AP) have weak mutual dependence of level 23%, 20.4%, 17.1% and 14.2% respectively. In addition to the indices $\hat{I}$ and $\hat{I}^*$, Table 9.1a presents the corresponding sample Kendall’s $\tau$ correlation coefficients. We observe that $\hat{I}^*$ and $\hat{I}$ recognize an association between two biomarkers almost in a similar way but, in general, the index $\hat{I}^*$ indicates a stronger dependence level. Notice that Kendall’s $\tau$ correlation coefficient cannot be used beyond the bivariate case.

The strong dependence between AST and ALT can be explained from the physiological point of view, since they are both released when hepatocytes undergo cell death. On the other hand, the weak dependence among the pairs of biomarkers can be explained by the fact that there are released under different mechanisms. When evaluating three biomarkers simultaneously, we
identify the strongest dependency among the triplet \((DB, AST, ALT)\). This observation can be explained from the physiologic point of view because processes that cause hepatocyte injury would also have a likely predilection to affect hepatocyte synthetic ability. This results in the inability of hepatocyte to conjugate bilirubin, as assessed by DB. The dependence among the different biomarkers we observe can, therefore, be explained in the context of liver function and injury.

In conclusion, we say that the high level of the mutual dependence of the random vector of the biomarkers is mainly due to the very strong dependence between biomarkers AST and ALT, and secondly, to a small extent, on the dependence between the biomarkers bilirubin and AST.

Table 9.1: The estimators of the indices \(\hat{I}_n\) and \(\hat{I}_n^*\) for the 2- and 3-dimensional sub-vectors of the vector of the biomarkers DB, AST, ALT and AP.

| Sub-vector | \(\hat{I}_n\) | \(\hat{I}_n^*\) | \(\hat{\tau}\) |
|------------|---------------|-----------------|-------------|
| (DB, AST)  | 0.176         | 0.112           | 0.099       |
| (DB, ALT)  | 0.354         | 0.230           | 0.204       |
| (DB, AP)   | 0.468         | 0.083           | 0.061       |
| (AST, ALT) | 0.792         | 0.171           | 0.142       |
| (AST, AP)  | 0.171         | 0.097           | 0.077       |
| (ALT, AP)  | 0.619         | 0.097           | 0.077       |

In addition to the preceding numerical results, Figures 9.1 and 9.2 shows the scatterplots of the bivariate and trivariate sub-vectors of the vector of biomarkers \((DB, AST, ALT, AP)\) respectively, illustrating graphically the dependence structure of these sub-vectors. Notice that even in the simplest case of bivariate random vectors it seems difficult to evaluate the strength of the dependence structure via scatterplots.

![Figure 9.1: The scatterplots of the bivariate sub-vectors of (DB, AST, ALT, AP).](image)
10. Concluding Remarks and Discussion

In this article we construct indices that measure joint or mutual dependence in a \( d \)-dimensional random vector. The indices depend on the Kendall process and satisfy the axioms set by Rényi that a function must satisfy in order to be called a measure of dependence. We evaluate the performance of the indices via simulation for a variety of distributions and sample sizes and present algorithms for their computation. Our proposed indices are copula-based and they are able to capture the degree of dependence exhibited in the data.

We also present a practical and fast, easy to compute, test for independence based on the estimated area under the Kendall curve, and propose an algorithm to compute quantities associated with its distribution. The performance of the proposed test statistic is evaluated via Monte Carlo simulation and is compared with the performance of the dHSIC-based test in terms of power and level, in a wide range of distributions and sample sizes. Our proposed test either outperforms or is competitive with the dHSIC-based test in most of the cases studied.

A. Lists of axioms for dependence measures

According to Schweizer and Wolff (1981), we can present Rényi’s conditions regarding a measure of dependence \( R(X,Y) \) for two continuously distributed variables \( X \) and \( Y \) in the following form.

\[ \text{A}_1^R: R(X,Y) \text{ is defined for any } X \text{ and } Y. \]
\[ \text{A}_2^R: R(X,Y) = R(Y,X). \]
\[ \text{A}_3^R: 0 \leq R(X,Y) \leq 1 \]
\[ \text{A}_4^R: R(X,Y) = 0 \text{ if and only if } X \text{ and } Y \text{ are independent.} \]
\[ \text{A}_5^R: R(X,Y) = 1 \text{ if and only if each of } X, Y \text{ is a.s. a strictly monotone function of the other.} \]
\[ \text{A}_6^R: \text{If } f \text{ and } g \text{ are strictly monotone a.s. on Range } X \text{ and Range } Y, \text{ respectively, then } R(f(X), g(Y)) = R(X,Y). \]
\[ \text{A}_7^R: \text{If the joint distribution of } X \text{ and } Y \text{ is bivariate normal, with correlation coefficient } \rho, \text{ then } R(X,Y) \text{ is a strictly increasing function of } |\rho|. \]
\[ \text{A}_8^R: \text{If } (X,Y) \text{ and } (X_n, Y_n), n = 1,2, \ldots, \text{ are pairs of random variables with joint distributions } H \text{ and } H_n, \text{ respectively, and if the sequence } \{H_n\} \text{ converges weakly to } H, \text{ then } \lim_{n \rightarrow \infty} R(X_n, Y_n) = R(X,Y). \]

Móri and Székely (2019) propose four natural axioms for dependence measures as follows. Let \( S \) be a nonempty set of pairs of nondegenerate random variables \( X, Y \) taking values in
Euclidean spaces or in real, separable Hilbert spaces $\mathcal{H}$. (Nondegenerate means that the random variable is not constant with probability 1.) Then $\mathcal{A}: \mathcal{S} \rightarrow [0, 1]$ is called a dependence measure on $\mathcal{S}$ if the following four axioms hold. In the axioms below we need similarity transformations of $\mathcal{H}$. Similarity of $\mathcal{H}$ is defined as a bijection (1–1 correspondence) from $\mathcal{H}$ onto itself that multiplies all distances by the same positive real number (scale). Similarities are known to be compositions of a translation, an orthogonal linear mapping, and a uniform scaling. They assume that if $(X, Y) \in \mathcal{S}$ then $(LX, MY) \in \mathcal{S}$ for all similarity transformations $L, M$ of $\mathcal{H}$.

$A_{1}^{MS}$: $\mathcal{A}(X, Y) = 0$ if and only if $X$ and $Y$ are independent.

$A_{2}^{MS}$: $\mathcal{A}(X, Y)$ is invariant with respect to all similarity transformations of $\mathcal{H}$; that is, $\mathcal{A}(LX, MY) = \mathcal{A}(X, Y)$ where $L, M$ are similarity transformations of $\mathcal{H}$.

$A_{3}^{MS}$: $\mathcal{A}(X, Y) = 1$ if and only if $Y = LX$ with probability 1, where $L$ is a similarity transformation of $\mathcal{H}$.

$A_{4}^{MS}$: $\mathcal{A}(X, Y)$ is continuous; that is, if $(X_n, Y_n) \in \mathcal{S}, n = 1, 2, \ldots$, such that for some positive constant $K$ we have $E(|X_n|^2 + |Y_n|^2) \leq K$, $n = 1, 2, \ldots$, and $(X_n, Y_n)$ converges weakly (converges in distribution) to $(X, Y)$ then $\mathcal{A}(X_n, Y_n) \rightarrow \mathcal{A}(X, Y)$. (The condition on the boundedness of second moments can be replaced by any other condition that guarantees the convergence of expectations: $E(X_n) \rightarrow E(X)$ and $E(Y_n) \rightarrow E(Y)$; such a condition is the uniform integrability of $X_n, Y_n$ which follows from the boundedness of second moments.)

## B. Kendall process

First, we state a modified empirical Kendall cdf as it is introduced by Genest and Rivest (1993), (see also Barbe et al., 1996). Let $\tilde{T}_{ni} = \{# j \neq i: X_j \leq X_i \text{ componentwise}\}/(n - 1)$. Denote by $\hat{F}_n$ the empirical distribution function of the $\tilde{T}_{ni}$'s that is an empirical Kendall cdf, i.e. an estimator of the Kendall cdf $K$.

We state now the following hypotheses.

$H_1$: The Kendall cdf $K(t)$ admits a continuous probability density function $k(t)$ on $[0, 1]$ that verifies $k(t) = o(t^{-1/2} \ln^{1/2-\varepsilon}(1/t))$ for some $\varepsilon > 0$ as $t \rightarrow 0$.

$H_2$: There exists a version of the conditional distribution of the vector $(F_1(X_1), \ldots, F_d(X_d))$ given $T = t$ and a countable family $\mathcal{P}$ of partitions $C$ of $[0, 1]^d$ into a finite number of Borel sets satisfying

$$\inf_{C \in \mathcal{P}} \max_{C \in C} \text{diam}(C) = 0,$$

such that for all $C \in \mathcal{C}$, the mapping $t \mapsto \mu_t(C) = k(t) \mathbb{P}\{(F_1(X_1), \ldots, F_d(X_d)) \in C | T = t\}$ is continuous on $(0, 1]$ with $\mu_1(C) = k(1) \mathbb{1}_{(1,...,1)}(C)$.

Given $0 \leq s, t \leq 1$, define

$$Q(s, t) = \mathbb{P}\{F(X_1) \leq s, X_1 \leq X_2 | F(X_2) = t\} = tK(s)$$

and

$$R(s, t) = \mathbb{P}\{X_1 \leq X_2 \wedge X_3 | F(X_2) = s, H(X_3) = t\} = st,$$

where $X_1, \ldots, X_3$ are independent and identically distributed random vectors from $F$, and $u \wedge v$ denotes the componentwise minimum between $u$ and $v$. 

Consider now the Kendall process $\mathbb{G}_n(t) \doteq \sqrt{n}\{\tilde{K}_n(t) - K(t)\}$, $0 \leq t \leq 1$. The following result is given by Barbe et al. (1996, Theo. 1), (cf. Genest and Rivest, 1993; Wellner, 2005; Salvadori et al., 2013).

**Theorem B.1.** Under $H_1$ and $H_2$, the Kendall process $\mathbb{G}_n(t) \doteq \sqrt{n}\{\tilde{K}_n(t) - K(t)\}$, $0 \leq t \leq 1$ converges in distribution to a Gaussian process $\mathbb{G}_K$ with zero mean and covariance function

$$\Gamma(s, t) = K(s \wedge t) - K(s)K(t) + k(s)k(t)R(s, t) - k(t)Q(s, t) - k(s)Q(t, s).$$

We restate here the hypothesis test $H_0 : F = \prod_{j=1}^d F_j$ versus $H_1 : F \neq \prod_{j=1}^d F_j$. Under $H_0$, $K_{H_0} = K_H$ and $k_{H_0} = k_H$; additionally, (see Barbe et al., 1996, Example 1, p. 202),

$$\sqrt{n}\{\tilde{K}_n(t) - K_H(t)\} \xrightarrow{d} \mathbb{G}_H,$$

(B.1)

where $\mathbb{G}_H$ is a Gaussian process with zero mean and for $0 \leq s, t \leq 1$ covariance function as it given by (7.2). In (7.2b) we have corrected a minor misprint (see Barbe et al., 1996, p. 208, for the correct general formula). Relations (7.2) lead to a cumbersome but explicit expression for the covariance function $\Gamma_H$, which reduces to $\tilde{\Gamma}_H(s, t) = s(t - 1 - \ln t)$, $0 \leq s \leq t \leq 1$, when $d = 2$.

A $\tilde{\mathbb{K}}_n$-based estimator of AUK is $\tilde{\text{AUK}}_n = E\{1 - K_H(\tilde{\tau}_n)\} = 1 - n^{-1}\sum_{i=1}^n K_H(\tilde{\tau}_i)$. Hence, using functional delta-method, we get $\sqrt{n}(\tilde{\text{AUK}}_n - 1/2) \xrightarrow{d} N(0, \sigma^2_H)$, where $\sigma^2_H$ is defined by (7.1b). When $d = 2$, after a straightforward calculation, we compute $\sigma^2_H = 19/423$. For completing the proof of (7.1) it is sufficient to prove that $\sqrt{n}(\tilde{\text{AUK}}_n - \text{AUK}_n) = o_p(1)$ as $n \to \infty$. For the proof of the asymptotic normality of both AUK and AUK see Appendix D.

### C. Useful results

The next result is the multivariate version of Pólya’s Theorem (Shao, 2003, Prop. 1.16, p. 51). First, denote by $\|f\|_\infty = \sup_{x \in \mathbb{R}^d} |f(x)|$ the supremum norm of $f : \mathbb{R}^d \to \mathbb{R}$.

**Theorem C.1.** Let $F$ and $F_n$ are cdfs on $\mathbb{R}^d$. If $F_n \Rightarrow F$ and $F$ is a continuous cdf, then $\|F_n - F\|_\infty \to 0$.

**Theorem C.2.** Let $X \sim F$ and $X_n \sim F_n$ on $\mathbb{R}^d$ such that $X_n \Rightarrow X$. Consider the random variables $T \doteq F(X)$ and $T_n \doteq F_n(X_n)$. If $F$ is continuous, $T_n \Rightarrow T$.

**Proof.** By Skorokhod’s representation Theorem, there are $\tilde{X}$ and $\tilde{X}_n$ defined in the same probability space $(\Omega, \mathcal{A}, \mathbb{P})$ such that $\tilde{X} \sim F$, $\tilde{X}_n \sim F_n$ and

$$\tilde{X}_n(\omega) \to \tilde{X}(\omega) \quad \text{for each } \omega \in \Omega,$$

(C.1)

(see, for example, Billingsley, 2013, Theo. 25.6). Consider now the random variables $\tilde{T} \doteq F(\tilde{X})$ and $\tilde{T}_n \doteq F_n(\tilde{X}_n)$; then, applying the triangle inequality, for each $\omega \in \Omega$,

$$|\tilde{T}_n(\omega) - \tilde{T}(\omega)| \leq |\tilde{F}_n(\tilde{X}_n(\omega)) - \tilde{F}(\tilde{X}_n(\omega))| + |\tilde{F}(\tilde{X}_n(\omega)) - \tilde{F}(\tilde{X}(\omega))|.$$

Because of continuity of $F$, (C.1) implies that

$$|\tilde{F}(\tilde{X}_n(\omega)) - \tilde{F}(\tilde{X}(\omega))| \to 0.$$
Using Theorem C.1, for each $\omega \in \Omega$,

$$|F_n(\bar{X}_n(\omega)) - F(\bar{X}_n(\omega))| \leq \|F_n - F\|_{\infty} \to 0.$$  

Consequently, $\bar{T}_n(\omega) \to \bar{T}(\omega)$ for each $\omega \in \Omega$; this implies that $\bar{T}_n \Rightarrow \bar{T}$ which gives $\bar{T}_n \Rightarrow T$, and so $T_n \Rightarrow T$. 

From Theorem C.2 and Faugeras (2013, Theo. 3.1, p. 274) the following theorem follows.

**Theorem C.3.** Let $X \sim F$ and $T$ as in Theorem C.1, and $(X_n)$ be an independent and identically distributed sequence from $F$. Consider the sequences of the ecdfs $\Xi_n$ as well as its empirical Kendall distribution functions $\Xi_n$. If $\bar{T}_n \sim \Xi_m$, then $\bar{T}_n \Rightarrow T$.

**D. Proofs of the main manuscript**

**Proof of Equation (4.2).** Suppose $Y$ is a continuous rv and consider the $d$-dimensional rvs $Y^{(1)} = (Y, \ldots, Y)$ and $Y^{(2)} = (Y, \ldots, -Y, -Y)T$ (the first $p \geq 1$ components of $Y^{(2)}$ are $Y$ and the last $d - p \geq 1$ are $-Y$). For each $y = (y_1, \ldots, y_d)^T \in \mathbb{R}^d$, one can easily see that $F_{Y^{(2)}}(y) = \mathbb{P}(Y \leq \min\{y_1, \ldots, y_d\})$; consequently, $T_{Y^{(1)}}(Y^{(1)}) = F_Y(Y) \sim U(0, 1)$, and so $K_{Y^{(1)}}(t) = t, t \in [0, 1]$. The AUK is

$$\text{AUK}_{Y^{(1)}} = \int_0^1 t \left(\frac{-\ln t}{(d-1)!}\right)^{d-1} dt = 2^{-d} \int_0^\infty \frac{2^d}{(d-1)!} x^{d-1} e^{-2x} dx = 2^{-d}.$$  

In the case of $Y^{(2)}$, we can easily to verify $F_{Y^{(2)}}(y) = \mathbb{P}(\max\{-y_{p+1}, \ldots, -y_d\} \leq Y \leq \min\{y_1, \ldots, y_p\}) = \max\{0, F_Y(\min\{y_1, \ldots, y_p\}) - F_Y(\max\{-y_{p+1}, \ldots, -y_d\})\}$; therefore, $T_{Y^{(2)}} = \max\{0, F_Y(Y) - F_Y(Y)\} = 0$ with probability one, and so $K_{Y^{(2)}}(t) = 1, t \in [0, 1]$. Thus,

$$\text{AUK}_{Y^{(2)}} = \mathbb{P}(T_{Y^{(2)}} \leq \Pi) = \mathbb{P}(0 \leq \Pi) = 1.$$  

For a continuous rv $X$, consider the $d$-dimensional $\bar{X} = (X, \ldots, X)^T$. The set $\mathcal{R}(\bar{X})$ contains 2 rvs of the kind of $Y^{(1)}$, specifically $(X, \ldots, X)^T$ and $(-X, \ldots, -X)^T$, and $2^d - 2$ of the kind of $Y^{(2)}$ (note that the coordinates of minus do not play any role). Hence, the corresponding vector $\bar{D}$ has 2 elements equals $2^{-d}$ and $2^d - 2$ elements equals 1; so, $\|D - A\| = (2^{d-2} - 2^{1-d} + 2^{1-2d})^{1/2}$, completing the proof.

**Proof of Proposition 5.1.** (a) By definition of $I$, $A_1$ holds in $\mathcal{X}_0^d$, while $A_2$ is obvious. Regarding $A_6$, for any such $f_i$'s define $s_i^l = 1$ if $f_i$ is increasing function and $-1$ when $f_i$ is decreasing function, and set $s^l = (s_1^l, \ldots, s_d^l)^T \in \mathbb{S}^d$, $\bar{\bar{X}} = \text{diag}(s^l) \tilde{X}$ and $f_{ij}(x_i) = f_i(s_i^lx_i), i = 1, \ldots, d$ (strictly increasing functions). Then, obviously, the random vectors $f(X) = (f_1(X_1), \ldots, f_d(X_d))^T$ and $\tilde{f}(X) = (\bar{f}_1(X_1), \ldots, \bar{f}_d(X_d))^T$ belong to $\mathcal{X}_0^d$ and $\mathcal{R}(f(X)) = \mathcal{R}(\tilde{f}(X))$. Observe now that for each $s_j = (s_{ij}, \ldots, s_{ij})^T \in \mathbb{S}^d$, $\text{diag}(s_j)f(X) = \text{diag}(s_j)\tilde{f}(\bar{\bar{X}})$, $\tilde{f}(\bar{\bar{X}}) = f^{(j)}(X) = f^{(j)}(x_i) = s_{ij}f_i(s_{ij}x_i)$ to be an strictly increasing function for each $i = 1, \ldots, d$. Hence, it becomes clear that it is enough to prove the desired result for strictly increasing functions; to prove this, one can use the similar arguments as in Vexler et al. (2019, proof of Proposition 4.4, p. 439). With regard to $A_2$, only for $d = 2$, it has been studied by Vexler et al. (2019). Finally, suppose $X_n \sim X \in \mathcal{X}_0^d$. Since $K_{II}(t), 0 \leq t \leq 1$, is a continuous and
bounded function, from Theorem C.2 (see Appendix C) and Portmanteau theorem we have that, for each \( j = 1, \ldots, 2^d \), \( \text{AUK}_{n_j} \to \text{AUK}_j \), and so, \( I(X_n) \to I(X) \); that, is \( A_8 \) holds.

(b) If \( X_1, \ldots, X_d \) are independent, one can easily see that \( \text{AUK}_j = 0 \) for all \( j = 1, \ldots, 2^d \), \( I = 0 \). Conversely, by definition, \( I = 0 \) implies that \( \text{AUK}_j = 0 \) for all \( j = 1, \ldots, 2^d \). Since \( X \in \mathcal{X}_d^1 \), there is at least an subscript \( j' \in \{1, \ldots, 2^d \} \) for which \( K_{j'}(t) = K(t) \) for \( 0 \leq t \leq 1 \). Hence, from Sklar’s Theorem (see Sklar, 1959; Durante et al., 2012; Salvadori et al., 2013) the corresponding (unique) copula to \( F_{X_{j'}} \) is the copula \( C_0(u) = u_1 \cdots u_d, u \in [0, 1]^d \), that corresponds to the independence case. Consequently, the components of \( X_{j'} \) are totally independent which is equivalent to the components of \( X \) are totally independent, i.e. \( F_X(x) = \prod_{i=1}^d F_{X_i}(x_i) \) for all \( x \in \mathbb{R}^d \).

(c) It is clear that \( I \geq 0 \) for all \( X \in \mathcal{X}_d^0 \). Suppose that \( X \in \mathcal{X}_d^1 \). Based on the Fréchet-Hoeffding upper-bound (see, for example, Rüschendorf, 2017), we have \( K(t) = \mathbb{P}(F(X) \leq t) \geq \mathbb{P}\left[ \min_{j=1,\ldots,d}\{F_j(X_j)\} \leq t \right] \geq \mathbb{P}\{F_1(X_1) \leq t\} = t \) for all \( 0 \leq t \leq 1 \); also, obviously, \( K(t) \leq 1 \) for all \( 0 \leq t \leq 1 \). Since \( X \in \mathcal{X}_d^1 \), there are at least two \( X_j, X_{j'} \in \mathcal{R}(X) \) for which \( C_2 \) holds; so, see in the proof of (4.2) above, \( \sum_{j=j',j''} \text{AUK}_{j''-1/2}^2 \leq 2(1-2^{-d})^2 \). On the other hand, it is obvious that \( \sum_{j=j',j''} \text{AUK}_{j''-1/2}^2 \leq (2^d - 2)/4 \). Hence, \( \sum_{j} \text{AUK}_{j-1/2}^2 \leq 2^{d-2} - 2^{d-1} + 2^{1-d} \), which implies that \( I \leq 1 \). Suppose now \( X \in \mathcal{X}_d^2 \) with \( I = 1 \). Then, there is a subscript \( j' \) such that \( K_{j'}(t) = t, 0 \leq t \leq 1 \). As in (b), from Sklar’s Theorem the corresponding (unique) copula to \( F_{X_{j'}} \) is the copula \( C_1(u) = \min\{u_1, \ldots, u_d\}, u \in [0, 1]^d \), that corresponds to the total dependence case.

\[ \square \]

Proof of Proposition 5.2. (a) See Proposition 5.1b. (b) See Proposition 5.1a. (c) Similar to Proposition 5.1c. \[ \square \]

Proof of Proposition 6.2. Suffice it to prove the result for the original (non-rotated) data. Set \( D_n = \|\mathbb{P}_n - F\|_\infty \). The multivariate version of Dvoretzky-Kiefer-Wolfowitz inequality says that (see Kiefer, 1961, Eq. (1.2), p. 649), there are positive constants \( c'_d \) and \( c_d \) (do not depend on \( F \)) such that, \( \mathbb{P}(D_n > z) \leq c'_d \exp(-c_d nz^2) \) for all \( n > 0 \) and \( z \geq 0 \). For \( \epsilon > 0 \), set \( \theta = \exp(-c_d \epsilon^2) \in (0, 1) \); then,

\[ \mathbb{P}\left( \sup_{m \geq n} D_m \geq \epsilon \right) \leq \sum_{m=n}^\infty \mathbb{P}(D_m > \epsilon) \leq c'_d \sum_{m=n}^\infty \theta^m = \frac{c'_d}{1 - \theta} \theta^m \to 0, \text{ as } n \to \infty. \]

Therefore, the multivariate version of Glivenko-Cantelli Theorem follows, that is,

\[ \|\mathbb{P}_n - F\|_\infty \overset{\text{a.s.}}{\to} 0. \]

Consider \( \xi(x) = K_{\mathbb{P}}\{F(x)\} \) and \( \hat{\xi}_n(x) = K_{\mathbb{P}_n}\{\mathbb{P}_n(x)\} \). Since \( K_{\mathbb{P}} \) is a continuous and bounded function, from the preceding result we have that \( \hat{\xi}_n \) convergence almost surely to \( \xi \) (uniformly). Let \( P \) and \( \mathbb{P}_n \) denote the probability measures that correspond to \( F \) and \( \mathbb{P}_n \) respectively. Because of bounding of \( \xi \), \( \text{AUK} = \int \hat{\xi}_n \, d\mathbb{P}_n \overset{\text{a.s.}}{\to} \int \xi \, dP = \text{AUK} \) (see van der Vaart, 1998, Sec. 19.4). \[ \square \]

Proof of the asymptotic normality of AUK and \( \hat{\text{AUK}} \). Consider the map \( K \mapsto \psi(K) \) defined by

\[ \psi(K) = \int_0^1 \{1 - K_{\mathbb{P}}(t)\} dK(t). \]
Observe that AUK = ψ(K) and AUK = ψ(Ł₀). Since AUKₖ = ψ(Kₖ) = 1/2, using (B.1), an application of functional delta-method (see van der Vaart, 1998, Theo. 20.8, p. 297) implies that

\[ \sqrt{n}(AUKₙ - 1/2) \overset{H₀}{\rightarrow} \psi'_{\text{H₁}}(G_{\text{H₁}}), \]

where \( \psi'_{\text{H₁}}(H) = \int_0^1 k_{\text{H₁}}(t)H(t)\,dt \) is the Hadamard derivative of \( \psi \) at \( K_{\text{H₁}} \). Due to the linearity of \( \psi'_{\text{H₁}}(\cdot) \), it follows that \( \psi'_{\text{H₁}}(G_{\text{H₁}}) \overset{d}{=} N(0, \sigma_{\text{H₁}}^2) \).

We now need to prove that \( \sqrt{n}(AUKₙ - AUK_n) = o_p(1) \) as \( n \to \infty \). The Kendall ecdf \( \hat{\mathbb{I}}_n \) is the ecdf of the values \( \hat{T}_{ni} = \{\#j : X_j \leq X_i \text{ componentwise} \}/n \); so AUKₙ = 1 - n⁻¹ \( \sum_{i=1}^n K_{\text{H₁}}(\hat{T}_{ni}) \) while AUKₙ = 1 - n⁻¹ \( \sum_{i=1}^n K_{\text{H₁}}(\hat{T}_{ni}) \). Therefore, \( n[AUKₙ - AUKₙ] = |\sum_{i=1}^n K_{\text{H₁}}(\hat{T}_{ni}) - K_{\text{H₁}}(\hat{T}_{ni})| \).

By definition, one can see that \( \hat{T}_{ni} - \hat{T}_{ni} = (1 - \hat{T}_{ni})/n \in [0, 1/n] \); hence, \( \hat{T}_{ni} \leq \hat{T}_{ni} \leq \hat{T}_{ni} + 1/n \). Suppose \( g(x) = x - x \ln x, x \geq 0 \), with \( g(0) = 0 \) and for each fixed \( \varepsilon > 0 \) consider the function \( h_{\varepsilon}(x) = g(x+\varepsilon) - g(x+\varepsilon), x \geq 0 \). Then, \( h_{\varepsilon} \) is a continuous function and \( h_{\varepsilon}'(x) = \ln(x/(x+\varepsilon)) < 0 \) for all \( x > 0 \). Hence, \( \max_{x \in [0, 1]} h_{\varepsilon}(x) = h_{\varepsilon}(0) = \varepsilon - \varepsilon \ln \varepsilon \). In view of the preceding results,

\[
\sum_{i=1}^n K_{\text{H₁}}(\hat{T}_{ni}) \leq \sum_{i=1}^n K_{\text{H₁}}(\hat{T}_{ni}) \leq \sum_{i=1}^n K_{\text{H₁}}(\hat{T}_{ni} + 1/n) \leq \sum_{i=1}^n K_{\text{H₁}}(\hat{T}_{ni}) + (1 + \ln n).
\]

From this immediately follows that AUKₙ ≤ AUKₙ and AUKₙ ≤ AUKₙ + (1 + ln n)/n. Thus,

\[ 0 \leq \sqrt{n}(AUKₙ - AUKₙ) \leq (1 + \ln n)/\sqrt{n}; \]

and so, \( \sqrt{n}(AUKₙ - AUKₙ) = o_p(1) \) as \( n \to \infty \) completing the proof of (7.1).

Suppose \( d = 2 \). It remains to compute the exact variance \( \sigma_{\text{H₁}}^2 \) to this case. First we present some preliminary results. Let \( k, n \) be nonnegative integers and \( t \in [0, 1] \). Define the integral

\[ I_{k,n}(t) = \int_0^t s^k \ln^n(1/s)\,ds. \]

Then, one can easily verify that

\[ I_{k,n}(t) = t^{k+1} \sum_{j=0}^n \frac{(n)_j}{(k+1)_j} \ln^{n-j}(1/t), \quad \text{with} \quad I_{k,n}(1) = \frac{n!}{(k+1)^n}, \]

where \((x)_j = x \cdots (x - j + 1)\) with \((x)_0 = 1\) denotes the \(j\)th order descending factorial function at \(x\). Finally, we compute

\[ \sigma_{\text{H₁}}^2 = 2 \int_0^1 \left\{ t \ln(1/t) - \ln(1/t) + \ln^2(1/t) \right\} \left\{ \int_0^t s \ln(1/s)\,ds \right\} \,dt, \]

that is

\[ \sigma_{\text{H₁}}^2 = 2 \int_0^1 \left\{ t \ln(1/t) - \ln(1/t) + \ln^2(1/t) \right\} I_{1,1}(t)\,dt = \frac{I_{1,1}}{2} + I_{3,2} - \frac{I_{2,1}}{2} - \frac{I_{2,2}}{2} + I_{2,3} = \frac{19}{432}. \]
Supplementary Material

**Title:** Supplementary Material: An AUK-based index for measuring and testing the joint dependence of a random vector

I. F-G-M copulas: Description of the Farlie–Gumbel–Morgenstern copulas that used in Subsubsection 8.1.2.

II. Bivariate distributions that are used in Subsection 8.2: Description of the bivariate distributions that are used in Subsection 8.2

III. Numerical results: The true values of the index \( I \) in the normal distribution case \( N(\mathbf{0}, \Sigma_3(\rho)) \) for various values of \( \rho \).

IV. A data set of biomarkers of hepatic injury: The data set of biomarkers of hepatic injury that used in Section 9.

V. R Codes: R codes to implement the developed method in this article.

Acknowledgements

The second author acknowledges financial support in the form of a grant award (award number 82114), from the Troup Fund, KALEIDA Health Foundation. The authors thank Dr. Andrew H. Talal for providing the data set and for helpful discussions aided the interpretation of the results.

References

Barbe, P., Genest, C., Ghoudi, K. and Rémillard, B. (1996). On Kendall’s process. *Journal of Multivariate Analysis*, 58, 197–229.

Bergsma, W. and Dassios, A. (2014). A consistent test of independence based on a sign covariance related to Kendall’s tau. *Bernoulli*, 20(2), 1006–1028.

Billingsley, P. (2013). *Convergence of probability measures*. John Wiley & Sons.

Brechmann, E. and Schepsmeier, U. (2013). Cdvine: Modeling dependence with c-and d-vine copulas in R. *Journal of Statistical Software*, 52, 1–27.

Durante, F., Fernández-Sánchez, J. and Sempi, C. (2012). Sklar’s theorem obtained via regularization techniques. *Nonlinear Analysis: Theory, Methods & Applications*, 75, 769–774.

Faugeras, O.P. (2013). Sklar’s theorem derived using probabilistic continuation and two consistency results. *Journal of Multivariate Analysis*, 122, 271–277.

Genest, C., Nešlehová, J.G., Rémillard, B. and Murphy, O.A. (2019). Testing for independence in arbitrary distributions. *Biometrika*, 106(1), 47–68.

Genest, C. and Rivest, L.-P. (1993). Statistical inference procedures for bivariate Archimedean copulas. *Journal of the American Statistical Association*, 88, 1034–1043.

Georges, P., Lamy, A.-G., Nicolas, E., Quibel, G. and Roncalli, T. (2001). Multivariate survival modelling: a unified approach with copulas. *Available at SSRN 1032559*.

Hofert, M., Oldford, W., Prasad, A. and Zhu, M. (2019). A framework for measuring association of random vectors via collapsed random variables. *Journal of Multivariate Analysis*, 172, 5–27.
JOE, H. (1997). *Multivariate models and multivariate dependence concepts*. Chapman and Hall/CRC.

JOHNSON, N.L. and KOTZ, S. (1975). On some generalized farlie-gumbel-morgenstern distributions. *Communications in Statistics-Theory and Methods*, **4**, 415–427.

KANKAINEN, A. (1995). Consistent testing of total independence based on the empirical characteristic function.

KARVANEN, J. (2005). A resampling test for the total independence of stationary time series: Application to the performance evaluation of ica algorithms. *Neural Processing Letters*, **22**(3), 311–324.

KENDALL, M.G. (1938). A new measure of rank correlation. *Biometrika*, **30**, 81–93.

KIIFER, J. (1961). On large deviations of the empiric df of vector chance variables and a law of the iterated logarithm. *Pacific Journal of Mathematics*, **11**, 649–660.

MÖR, T.F. and SZÉKELY, G.J. (2019). Four simple axioms of dependence measures. *Metrika*, **82**, 1–16.

PEARSON, K. (1895). Note on regression and inheritance in the case of two parents. *Proceedings of the Royal Society of London*, **58**, 240–242.

PFISTER, N., BÜHLMANN, P., SCHÖLCKOPF, B. and PETERS, J. (2018). Kernel-based tests for joint independence. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **80**(1), 5–31.

PFISTER, N. and PETERS, J. (2019). Package ‘dHSIC’. R package version 2.1. https://CRAN.R-project.org/package=dHSIC

RéNYI, A. (1959). On measures of dependence. *Acta Mathematica Hungarica*, **10**, 441–451.

RÜSCHENDORF, L. (2017). Improved Hoeffding–Fréchet bounds and applications to VaR estimates. In *Copulas and Dependence Models with Applications* (pp. 181-202). Cham: Springer International Publishing.

SALVADORI, G., DURANTE, F. and PERRONE, E. (2013). Semi-parametric approximation of Kendall’s distribution function and multivariate return periods. *Journal de la Société Française de Statistique*, **154**, 151–173.

SCHWEIZER, B. and WOLF, E.F. (1981). On nonparametric measures of dependence for random variables. *The Annals of Statistics*, **9**, 879–885.

SHEN, C., PIEBRE, C.E. and VOGELSTEIN, J.T. (2020). From distance correlation to multiscale graph correlation. *Journal of the American Statistical Association*, **115**(529), 280–291.

SHAO, J. (2003). *Mathematical statistics*. Springer Texts in Statistics (2nd ed.). Springer-Verlag, New York.

SIMON, G. (1977). A nonparametric test of total independence based on Kendall’s tau. *Biometrika*, **64**(2), 277–282.

SKLAR, M. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publications de l’Institut de statistique de l’Université de Paris*, **8**, 229–231.

SPEARMAN, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychology*, **15**, 72–101.

SZÉKELY, G.J. and RIZZO, M.L. (2009). Brownian distance covariance (with discussion). *The Annals of Applied Statistics*, 1236–1368.

SZÉKELY, G.J. and RIZZO, M.L. (2014). Partial distance correlation with methods for dissimilarities. *The Annals of Statistics*, **42**(6), 2382–2412.

SZÉKELY, G.J., RIZZO, M.L. and BAKIROV, N.K. (2007). Measuring and testing dependence by
correlation of distances. *The Annals of Statistics, 35*(6), 2769–2794.

Turlach, B.A. and Murray, K. (2019). Package ‘MonoPoly’. https://CRAN.R-project.org/package=MonoPoly

Vand der Vaart, A.W. (1998). *Asymptotic statistics* (volume 3). Cambridge University Press, Cambridge.

Vexler, A., Afendras, G. and Markatou, M. (2019). Multi-panel Kendall plot in light of an roc curve analysis applied to measuring dependence. *Statistics, 53*, 417–439.

Wellner, J.A. (2005). Empirical processes: Theory and Applications. *Notes for a course given at Delft University of Technology.*

Zhu, L., Xu, K., Li, R. and Zhong, W. (2017). Projection correlation between two random vectors. *Biometrika, 104*(4), 829–843.
**Supplementary Material:** An AUK-based Index for Measuring and Testing the Joint Dependence of a Random Vector

Georgios Afendras

Department of Mathematics, Aristotle University of Thessaloniki, Thessaloniki, Greece

Marianthi Markatou

Department of Biostatistics and Jacobs School of Medicine and Biomedical Sciences, The State University of New York at Buffalo, Buffalo, NY, USA

Albert Vexler

Department of Biostatistics, The State University of New York at Buffalo, Buffalo, NY, USA

I. F-G-M copulas

We consider next two special cases of trivariate F-G-M copulas. For \( u = (u_1, u_2, u_3)^T \in [0, 1]^3 \), consider the copulas

\[
C_\theta(u) = u_1 u_2 u_3 \{1 + \theta(1 - u_1)(1 - u_2)\};
\]

\[
\tilde{C}_\theta(u) = u_1 u_2 u_3 \{1 + \theta(1 - u_1)(1 - u_2)(1 - u_3)\},
\]

where \( |\theta| \leq 1 \) (see Johnson and Kotz, 1975; Georges et al., 2001, Remark 4, p. 9). Obviously, both \( C_\theta=0 \) and \( \tilde{C}_\theta=0 \) correspond to the independent copula case \( u_1 u_2 u_3 \).

Suppose \( U = (U_1, U_2, U_3)^T \sim C_\theta \), see (I.1a). Obviously, \( (U_1, U_2)^T \) and \( U_3 \) are independent; also, one can easily verify that the cdf of \( U_1|U_2 = u_2 \) is \( F_{U_1|U_2}(u_1|u_2) = \alpha u_1^2 + (1 - \alpha) u_1, 0 \leq u_1 \leq 1 \) (for each \( 0 \leq u_2 \leq 1 \)), where \( \alpha \equiv \alpha(u_2; \theta) = \theta(2u_2 - 1) \). Thus,

\[
F_{U_1|U_2}(u_1|u_2) = \begin{cases} u_1, & \text{if } \alpha = 0, \\ u_1,1 \mathbb{I}_{\{0 \leq u_1,1 \leq 1\}} + u_1,2 \mathbb{I}_{\{0 \leq u_{1,2} \leq 1\}}, & \text{if } \alpha \neq 0, \end{cases}
\]

where \( u_{1,1} = [\alpha - 1 - \{(1 - \alpha)^2 + 4\alpha u_1\}^{1/2}] / (2\alpha) \) and \( u_{1,2} = [\alpha - 1 + \{(1 - \alpha)^2 + 4\alpha u_1\}^{1/2}] / (2\alpha) \).

Let now \( \tilde{U} = (\tilde{U}_1, \tilde{U}_2, \tilde{U}_3)^T \sim \tilde{C}_\theta \), see (I.1b). The cdf of \( (\tilde{U}_2, \tilde{U}_3)^T \) is \( F_{\tilde{U}_2,\tilde{U}_3}(u_2, u_3) = u_2 u_3, (u_2, u_3)^T \in [0, 1]^2 \), that is \( \tilde{U}_2 \) and \( \tilde{U}_3 \) are independent. Moreover, the cdf of \( \tilde{U}_1|\tilde{U}_2, \tilde{U}_3 = (u_2, u_3)^T \) is \( F_{\tilde{U}_1|\tilde{U}_2,\tilde{U}_3}(u_1|u_2, u_3) = \tilde{u} u_1^2 + (1 - \tilde{u}) u_1, 0 \leq u_1 \leq 1 \) (for each \( 0 \leq u_2, u_3 \leq 1 \)), where \( \tilde{\alpha} \equiv \tilde{\alpha}(u_2, u_3; \theta) = -\theta(2u_2 - 1)(2u_3 - 1) \). Thus,

\[
\tilde{F}^{-1}_{\tilde{U}_1|\tilde{U}_2,\tilde{U}_3}(u_1|u_2, u_3) = \begin{cases} u_1, & \text{if } \tilde{\alpha} = 0, \\ \tilde{u}_{1,1} \mathbb{I}_{\{0 \leq \tilde{u}_{1,1} \leq 1\}} + \tilde{u}_{1,2} \mathbb{I}_{\{0 \leq \tilde{u}_{1,2} \leq 1\}}, & \text{if } \tilde{\alpha} \neq 0, \end{cases}
\]

where \( \tilde{u}_{1,1} = [\tilde{\alpha} - 1 - \{(1 - \tilde{\alpha})^2 + 4\tilde{\alpha} u_1\}^{1/2}] / (2\tilde{\alpha}) \) and \( \tilde{u}_{1,2} = [\tilde{\alpha} - 1 + \{(1 - \tilde{\alpha})^2 + 4\tilde{\alpha} u_1\}^{1/2}] / (2\tilde{\alpha}) \).

From the preceding analysis, if \( X_1, X_2 \) and \( X_3 \) are independent \( U(0,1) \) distributions, then \( (F^{-1}_{U_1|U_2}(X_1|X_2), X_2, X_3)^T \sim C_\theta \) and \( (\tilde{F}^{-1}_{\tilde{U}_1|\tilde{U}_2,\tilde{U}_3}(X_1|X_2, X_3), X_2, X_3)^T \sim \tilde{C}_\theta \).
II. Bivariate distributions that are used in Subsection 8.2

Bivariate uniform distribution on the circumference of a circle, \( U\{C(0, 1)\} \) Let \( Z_1, Z_2 \) be iid standard normal distributions. Consider the rvs \( U_i = Z_i / (Z_1^2 + Z_2^2)^{1/2}, \ i = 1, 2. \) Then, \((U_1, U_2)\) is uniformly distributed on the circumference of the circle \( x^2 + y^2 = 1. \)

Bivariate exp\(\{a_1, a_2, a_{12}\}\) distribution Let \( a_1, a_2, a_{12} > 0 \) with \( a_{12} < \min\{a_1, a_2\}. \) Suppose that \( E_1 \sim \exp(a_1 - a_{12}), E_2 \sim \exp(a_2 - a_{12}) \) and \( E_3 \sim \exp(a_{12}), \) and consider the rvs \( X_i = \min\{E_i, E_3\} - 1/a_i, \ i = 1, 2. \) Then, \((X_1, X_2)\) follows the bivariate exponential distribution.

Bivariate Morgenstern\(\{a\}\) distribution Let \( a > 0 \) and \( X, U \) be independent \( U(0, 1) \) distributed. Consider the rvs \( Z = a(2X - 1) - 1, W = 1 - 2a(2X - 1) + a^2(2X - 1)^2 + 4aU(2X - 1) \) and \( Y = 2U/(W^{1/2} - Z). \) Then, \((X, Y)\) follows the bivariate Morgenstern distribution.

Bivariate Plackett\(\{s\}\) distribution Let \( s > 1 \) and \( X, U \) be independent \( U(0, 1) \) distributed. Consider the rvs \( W_1 = U(1 - U), W_2 = s + W_1(s - 1)^2, W_3 = 2W_1(s^2X + 1 - X) + s(1 - 2W_1), \) \( W_4 = s + 4(1 - s)^2X(1 - X)W_1 \) and \( Y = W_2\{W_3 - (1 - 2U)W_4^2\}/2. \) Then, \((X, Y)\) follows the bivariate Plackett distribution.

Bivariate ALI-HAQ’s\(\{a, p\}\) distribution Let \( a > 0, 0 < p < 1 \) and \( X, U \) be independent \( U(0, 1) \) distributed. Consider the rvs \( V = 1 - U, W_1 = a(2UV + 1) + 2a^2V^2U + 1, \) \( W_2 = a^2(4V^2U - 4VU + 1) + a(4VU - 4p + 2) + 1 \) and \( Y = 2U(aV - 1)/\sqrt{(W_1 + W_2^2)}. \) Then, \((X, Y)\) follows the bivariate ALI-HAQ’s distribution.

Bivariate Gumbel\(\{e\}\) distribution Let \( e > 0 \) and \( U_1, \ldots, U_4 \) be independent \( U(0, 1) \) distributed. Consider the rvs \( X = -\ln U_1, W_1 = 1 + eX, W_2 = (W_1 - e)/W_1, W_3 = -\ln U_2 \) and \( Y = 1_{\{U_3 < W_2\}}W_1W_3 + 1_{\{U_3 \geq W_2\}}W_1(W_3 - \ln U_4). \) Then, \((X, Y)\) follows the bivariate Gumbel distribution.

Bivariate \(t_5\) distribution Using the package “mvtnorm” in R, we simulate data from bivariate student distribution with 5 degree of freedom and variance-covariance matrix \( \Sigma = \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix}. \)
III. Numerical results

This section presents some useful numerical results. Table III.1 contains the true values of the index $I$ in the normal distribution case $N_3(0, \Sigma_3(\rho))$, where the variance/covariance matrix $\Sigma_3(\rho)$ is defined by (2.2b), for various values of $\rho$. Additionally, Figure III.1 presents scatterplots based on random samples from $N_3(0, \Sigma_3(\rho))$ distributions for various values of $\rho$, giving an illustration of the shapes of these distributions.

Table III.1: The index $I \equiv I(\rho)$ for various values of $\rho$ when the underline distribution is $N_3(0, \Sigma_3(\rho))$.

| $\rho$: | 0    | 0.05 | 0.10 | 0.15 | 0.20 | 0.25 | 0.30 | 0.35 | 0.40 | 0.45 | 0.50 | 0.55 |
|--------|------|------|------|------|------|------|------|------|------|------|------|------|
| $I$:   | 0.027| 0.052| 0.077| 0.101| 0.124| 0.148| 0.171| 0.194| 0.218| 0.243| 0.267|
| $\rho$: | 0.60 | 0.65 | 0.70 | 0.75 | 0.80 | 0.85 | 0.90 | 0.95 | 0.97 | 0.98 | 0.99 | 0.995|
| $I$:   | 0.294| 0.321| 0.352| 0.385| 0.423| 0.467| 0.524| 0.609| 0.662| 0.701| 0.758| 0.841|

Figure III.1: Scatterplots: Random samples of size $n = 1000$ are drawn from the trivariate normal distributions $N_3(0, \Sigma_3(\rho))$ for various values of $\rho$. 

(a) $\rho = 0$.  
(b) $\rho = 0.25$.  
(c) $\rho = 0.4$.  
(d) $\rho = 0.5$.  
(e) $\rho = 0.6$.  
(f) $\rho = 0.75$.  
(g) $\rho = 0.8$.  
(h) $\rho = 1$.  


## IV. A data set of biomarkers of hepatic injury

Table IV.1: A data set of size \( n = 208 \) of the biomarkers Bilirubin (B), Alanine and Aspartate aminotransferase (ALT and AST) and Alkaline phosphatase (AP).

| PID | B   | AST | ALT | AP   | PID | B   | AST | ALT | AP   | PID | B   | AST | ALT | AP   |
|-----|-----|-----|-----|------|-----|-----|-----|-----|------|-----|-----|-----|-----|------|
| 1   | 1.1 | 34  | 54  | 150  | 2   | 0.3 | 22  | 33  | 90   | 3   | 0.4 | 17  | 21  | 105  |
| 2   | 0.3 | 22  | 33  | 90   | 3   | 0.4 | 17  | 21  | 105  | 4   | 0.2 | 10  | 17  | 178  |
| 3   | 0.4 | 17  | 21  | 105  | 4   | 0.2 | 10  | 17  | 178  | 5   | 0.5 | 26  | 38  | 57   |
| 5   | 0.5 | 26  | 38  | 57   | 6   | 1.6 | 144 | 116 | 158  | 7   | 1.1 | 51  | 28  | 145  |
| 7   | 1.1 | 51  | 28  | 145  | 8   | 0.5 | 26  | 38  | 57   | 9   | 1.3 | 48  | 31  | 295  |
| 9   | 1.3 | 48  | 31  | 295  | 10  | 0.9 | 25  | 30  | 76   | 11  | 0.7 | 57  | 52  | 88   |
| 11  | 0.7 | 57  | 52  | 88   | 12  | 0.2 | 34  | 35  | 76   | 13  | 0.5 | 29  | 14  | 66   |
| 14  | 1.7 | 30  | 121 | 66   | 15  | 1.7 | 21  | 13  | 94   | 16  | 1.2 | 26  | 16  | 207  |
| 16  | 1.2 | 26  | 16  | 207  | 17  | 0.9 | 34  | 38  | 59   | 18  | 0.8 | 42  | 89  | 72   |
| 18  | 0.8 | 42  | 89  | 72   | 19  | 0.5 | 50  | 73  | 68   | 20  | 0.4 | 43  | 35  | 87   |
| 20  | 0.4 | 43  | 35  | 87   | 21  | 0.8 | 19  | 26  | 112  | 22  | 0.7 | 16  | 11  | 36   |
| 22  | 0.7 | 16  | 11  | 36   | 23  | 0.4 | 20  | 22  | 99   | 24  | 0.5 | 40  | 58  | 71   |
| 24  | 0.5 | 40  | 58  | 71   | 25  | 0.5 | 44  | 59  | 79   | 26  | 1.2 | 30  | 22  | 49   |
| 26  | 1.2 | 30  | 22  | 49   | 27  | 0.2 | 20  | 17  | 118  | 28  | 0.7 | 36  | 27  | 129  |
| 28  | 0.7 | 36  | 27  | 129  | 29  | 0.9 | 20  | 22  | 89   | 30  | 0.5 | 32  | 18  | 147  |
| 30  | 0.5 | 32  | 18  | 147  | 31  | 0.3 | 25  | 28  | 80   | 32  | 0.9 | 20  | 22  | 89   |
| 32  | 0.9 | 20  | 22  | 89   | 33  | 1.2 | 55  | 61  | 52   | 34  | 1.7 | 34  | 14  | 75   |
| 34  | 1.7 | 34  | 14  | 75   | 35  | 1.1 | 35  | 39  | 110  | 36  | 0.4 | 21  | 27  | 75   |
| 36  | 0.4 | 21  | 27  | 75   | 37  | 0.4 | 24  | 30  | 77   | 38  | 0.5 | 38  | 37  | 67   |
| 38  | 0.5 | 38  | 37  | 67   | 39  | 0.7 | 22  | 22  | 56   | 40  | 0.6 | 53  | 37  | 128  |
| 40  | 0.6 | 53  | 37  | 128  | 41  | 0.4 | 25  | 21  | 97   | 42  | 0.5 | 52  | 75  | 68   |
| 42  | 0.5 | 52  | 75  | 68   | 43  | 0.6 | 28  | 33  | 142  | 44  | 1  | 33  | 25  | 59   |
| 44  | 1   | 33  | 25  | 59   | 45  | 0.4 | 28  | 39  | 164  | 46  | 0.3 | 39  | 47  | 98   |
| 46  | 0.3 | 39  | 47  | 98   | 47  | 0.8 | 28  | 41  | 68   | 48  | 1.2 | 64  | 26  | 172  |
| 48  | 1.2 | 64  | 26  | 172  | 49  | 0.8 | 45  | 47  | 98   | 50  | 3.7 | 74  | 102 | 245  |
| 50  | 3.7 | 74  | 102 | 245  | 51  | 0.8 | 32  | 35  | 57   | 52  | 1 | 293 | 227 | 147  |
V. R codes

The required packages that used.

```r
rm(list=ls())
library(MASS); require(MASS); sfLibrary(MASS)
library(energy); require(energy)
library(Emcdf); sfLibrary(Emcdf)
library(mvtnorm)
library(gtools)
library/copula
library("snowfall")
library("rlecuyer")
sfInit(parallel=TRUE, cpus=8, type="SOCK")
library(dHSIC)
```

The specific formula of \( 1 - \Phi_H(t) \) when \( d = 3 \).

```r
g <- function(t){ ifelse(t==0,1,1-t*t*log(t)-0.5*t*(log(t))^2)}
g <- Vectorize(g)
```

The pdf \( f \) of \( N_3(0, \Xi_3(\rho)) \) distribution.

```r
f <- function(x,y,z,r12,r13,r23){
S <- matrix(c(1,r12,r13,r12,1,r23,r13,r23,1),3,3)
density <- dmvnorm(c(x,y,z), mean=c(0,0,0), sigma=S)
return(density)
} f <- Vectorize(f)
```

The cdf \( F \) of \( N_3(0, \Xi_3(\rho)) \) distribution.

```r
F <- function(x,y,z,r12,r13,r23){
S <- matrix(c(1,r12,r13,r12,1,r23,r13,r23,1),3,3)
Pr <- pmvnorm(lower=rep(-Inf,3), upper=c(x,y,z),
mean=c(0,0,0), corr=S)
return(Pr)
} F <- Vectorize(F)
```

Code for computing AUK when \( X \sim N_3(0, \Xi_3(\rho)) \).

```r
AUK <- function(r12,r13,r23){
auk <- integrate(Vectorize(function(x) {
integrate(Vectorize(function(y) {
integrate(function(z) {
g(F(x,y,z,r12,r13,r23))*f(x,y,z,r12,r13,r23)},
-Inf, Inf}$value )}, -Inf,Inf)$value )}, -Inf,Inf)$value
return(auk)}
```

Code for computing \( I \) when \( X \sim N_3(0, \Xi_3(\rho)) \).

```r
I <- function(rho){
Sign <- permutations(n=2,r=3,v=c(1,-1), repeats.allowed=TRUE)
auk <- rep(0,8)
```
for (j in 1:8){
    r12 = Sign[j,1]*Sign[j,2]*rho
    r13 = Sign[j,1]*Sign[j,3]*rho
    r23 = Sign[j,2]*Sign[j,3]*rho
    auk[j] <- AUK(r12, r13, r23)
} 

ind=sqrt(32/57)*sqrt(sum((auk-1/2)^2))
return(ind)}

THE ECDF $F_n$ OF A DATA SET $X_n = (X_1, \ldots, X_n)^T$.

Fn <- function(Xn){
    Fn1 <- function(i){emcdf(Xn, Xn[i,])}
    Fn1 <- Vectorize(Fn1)
    FFn <- Fn1(c(1:nrow(Xn)))
    return(FFn)}

THE AUK$_n$ ESTIMATOR BASED ON THE DATA SET $X_n = (X_1, \ldots, X_n)^T$.

AUK <- function(Xn){mean(g(Fn(Xn)))}

CODE FOR COMPUTING $I$, VIA MONTE CARLO, WHEN $X \sim N_3(0, \Sigma_3(\rho))$.

I <- function(N,n,rho){
    mu <- c(0,0,0)
    Sigma <- rho*matrix(rep(1,9),3,3)+(1-rho)*diag(3)
    I1 <- function(i){
        if (i==1){
            X <- mvnorm(n, mu, Sigma)
            Sign <- permutations(n=2, r=3, v=c(1,-1), repeats.allowed=TRUE)
            auk1 <- function(j){AUK(X%*%diag(Sign[j,]))}
            auk1 <- Vectorize(auk1)
            auk1 <- auk1(1:8)
            A=sqrt(32/57)*sqrt(sum((auk-1/2)^2))
            return(A)}
        Vectorize(I1)
    ind <- mean(I1(rep(1,N)))
    return(ind)
} 
I <- Vectorize(I)

THE FUNCTION $\varphi_3$.

phi <- function(t){1.61*t+4.513*t^2-13.607*t^3+12.235*t^4-3.751*t^5}
phi <- Vectorize(phi)

THE $\hat{I}_n$ AND $\hat{I}'_n$ ESTIMATORS BASED ON THE DATA SET $X_n = (X_1, \ldots, X_n)^T$.

Indices <- function(X){
    Sign <- permutations(n=2, r=3, v=c(1,-1), repeats.allowed=TRUE)
    auk1 <- function(j){AUK(X%*%diag(Sign[j,]))}
    auk1 <- Vectorize(auk1)
    auk <- auk1(1:8)
    ind=sqrt(32/57)*sqrt(sum((auk-1/2)^2))
    ind.st <- phi(ind)
    return(c(ind, ind.st))}
**ESTIMATING THE FUNCTION ϕ_d USING MonoPoly R PACKAGE.**

Let \( \varrho \) and \( I \) be as in Step 1 of the algorithm

\[
C \leftarrow \text{SOSpol.fit(1, rho, deg.is.odd=TRUE, K=3, a=0, b=1, monotone = c("increasing", "decreasing"), trace = FALSE, plot.it = FALSE, type=2, control = monpol.control())}
\]

\[
C \leftarrow C-C[1]
\]

\[
\phi_d \leftarrow \text{function(t){evalPol(t.C)}}
\]

\[
\phi_d \leftarrow \text{Vectorize(\phi.d)}
\]

**ESTIMATING THE STANDARD DEVIATION \( \sigma_H \).**

Wrapper <- function(d){
  \[
  l \leftarrow \text{length(d)}
  \]
  Fn <- function(Un){
    Fn1 <- function(i){
      \[
      \text{emcfdf(Un, Un[i,])}
      \]
    }
    Fn1 <- Vectorize(Fn1)
    FFn <- Fn1(c(1:nrow(Un))
    return(FFn)
  }
  sigma_Pi <- function(r, n, d){
    AUK1 <- function(i){
      if (i==1){
        \[
        X \leftarrow \text{matrix(runif(2*n, min=0, max=1), n, 2)}
        \]
        auk <- mean(pchisq(-2*log(Fn(X)), df=2*d, ncp=0, lower.tail=TRUE))
        return(auk)
      }
    }
    AUK1 <- Vectorize(AUK1)
    AUK <- AUK1(rep(1, r))
    Zn <- sqrt(n)*(AUK-1/2)
    sigmaPi <- sqrt(var(Zn))
    sigma_Pi <- Vectorize(sigma_Pi)
    return(sigmaPi)
  }
  sPi <- sigma_Pi(1000, 50000, d)
}

start_time <- Sys.time()
Output <- sflapply(Value.list, Wrapper)
R <- matrix(unlist(Output), nrow=length(d), ncol=1)
rownames(R) <- d
colnames(R) <- c("sigma_Pi")
end_time <- Sys.time()
running_time <- end_time - start_time
sfStop()
A code to measure the speed-time of AUK and dHSIC tests

```r
speed.times <- function(n) {
  x <- runif(n)
  y <- runif(n)
  X <- cbind(x, y)
  StartTime.AUK <- Sys.time()
  AUK <- mean(g(Fn(X)))
  Z <- sqrt(n) * abs(AUK - 1/2) / (sqrt(19/432))
  EndTime.AUK <- Sys.time()
  TimeTaken.AUK <- EndTime.AUK - StartTime.AUK
  StartTime.dHSIC <- Sys.time()
  dHSIC <- dhsic.test(list(x, y), alpha = 0.05, method = "permutation",
                      kernel = "gaussian", B = 1000, pairwise = FALSE,
                      bandwidth = 1, matrix.input = FALSE)
  EndTime.dHSIC <- Sys.time()
  TimeTaken.dHSIC <- EndTime.dHSIC - StartTime.dHSIC
  Res <- c(n, c(TimeTaken.AUK, TimeTaken.dHSIC))
  R <- matrix(Res, nrow=1, ncol=3)
  colnames(R) <- c("sample.size\(n\)", "AUK\-time\(\) (in sec)",
                  "dHSIC\-time\(\) (in sec)")
  rownames(R) <- c(""")
  return(t(R))
}
```