Most of the popular dependence measures for two random variables $X$ and $Y$ (such as Pearson’s and Spearman’s correlation, Kendall’s $\tau$ and Gini’s $\gamma$) vanish whenever $X$ and $Y$ are independent. However, neither does a vanishing dependence measure necessarily imply independence, nor does a measure equal to 1 imply that one variable is a measurable function of the other. Yet, both properties are natural properties for a convincing dependence measure.

In this paper, we present a general approach to transforming a given dependence measure into a new one which exactly characterizes independence as well as functional dependence. Our approach uses the concept of monotone rearrangements as introduced by Hardy and Littlewood and is applicable to a broad class of measures. In particular, we are able to define a rearranged Spearman’s $\rho$ and a rearranged Kendall’s $\tau$ which do attain the value 0 if and only if both variables are independent, and the value 1 if and only if one variable is a measurable function of the other. We also present simple estimators for the rearranged dependence measures, prove their consistency and illustrate their finite sample properties by means of a simulation study and a data example.

**MSC2020 subject classifications:** Primary 62H20; secondary 62H05

**Keywords:** measure of dependence; coefficient of correlation; decreasing rearrangement; copula

1. Introduction

One of the most fundamental problems in statistics is to measure the association between two random variables $X$ and $Y$ based on a sample of independent identically distributed observations $(X_1, Y_1), \ldots, (X_n, Y_n)$, and numerous proposals have been made for this purpose. These measures usually vary in the interval $[0, 1]$ or $[-1, 1]$, and vanish if the variables are independent. Moreover, many of these measures, including the frequently used Pearson’s and Spearman’s correlation, Kendall’s $\tau$ and Gini’s $\gamma$, are very powerful to detect linear and monotone dependencies. On the other hand, in general, a vanishing dependence measure (such as Pearson’s coefficient) only implies independence of $X$ and $Y$ under quite restrictive additional assumptions (such as a normal distribution), and it is a well known fact that many of these measures cannot detect non-monotone associations.

Several authors have proposed solutions to this problem by introducing alternative dependence measures, but mainly in the context of testing for independence. Among the many contributions, we mention exemplary the early work of Blum, Kiefer and Rosenblatt (1961), Csörgő (1985), Rosenblatt (1975), Schweizer and Wolff (1981) and the more recent papers by Bergsma and Dassios (2014), Gretton et al. (2008), Székely, Rizzo and Bakirov (2007) and Zhang (2019). However, as pointed out by Chatterjee (2021), these measures are designed primarily for testing independence, and not for measuring the strength of the relationship between the variables. In the same paper, a correlation coefficient is presented which estimates a (population) measure $\mu$ of the dependence between two random variables $X$ and $Y$ with the following properties:

\begin{equation}
0 \leq \mu(X, Y) \leq 1
\end{equation}
\[(1.2) \quad \mu(X, Y) = 0 \text{ if, and only if, } X \text{ and } Y \text{ are independent}\]

\[(1.3) \quad \mu(X, Y) = 1 \text{ if, and only if, } Y = f(X) \text{ for some measurable function } f : \mathbb{R} \to \mathbb{R}.\]

For continuous distributions the measure \(\mu\) has been introduced and studied in Dette, Siburg and Stoimenov (2013) who also proposed a kernel based estimator for it. Since its introduction, Chatterjee’s correlation coefficient has found considerable attention in the literature (see Audy, Deb and Nandy, 2021, Cao and Bickel, 2020, Deb, Ghosal and Sen, 2020, Gamboa et al., 2022, Lin and Han, 2022, Shi, Drton and Han, 2021, 2022, among others), which underlines the demand for dependence measures possessing the above properties (1.1)–(1.3).

This paper takes a quite different viewpoint on this problem by formulating the following question:

*Is it possible to transform a given dependence measure in such a way that the new dependence measure satisfies properties (1.1)–(1.3)?*

Our answer to this question is affirmative. More precisely, we will show that there exists a well defined transformation \(\mu \mapsto R_\mu\) with the following property. Whenever the dependence measure \(\mu\) satisfies the axioms (1.1) to (1.3) on the set of stochastically increasing continuous distributions, the new dependence measure \(R_\mu\) will satisfy (1.1) to (1.3) on the set of all continuous distributions. By definition, a pair \((X, Y)\) of random variables is stochastically increasing if the function \(x \mapsto \mathbb{P}(Y \leq y | X = x)\) is decreasing for each fixed \(y\) (see, e.g. Nelsen, 2006). This property was also discussed earlier in Lehmann (1959) under the term *positive regression dependence*.

The transformed dependence measure \(R_\mu\) will be called the *rearranged dependence measure*. It turns out that the new transformation is applicable to many of the classical dependence measures and, consequently, enables us to define rearranged dependence measures such as the rearranged Spearman’s \(\rho\) and the rearranged Kendall’s \(\tau\), all of which satisfy properties (1.1)–(1.3).

Our approach is based on a classical concept from majorization theory which is called *monotone rearrangement* (see, for instance, Hardy, Littlewood and Polya, 1988, Ryff, 1965, 1970). In the last decades, monotone rearrangements have found considerable interest in the statistical literature. For example, Anevski and Fougères (2019), Camirand-Lemyre, Carroll and Delaigle (2022), Chernozhukov, Fernández-Val and Galichon (2009), Dette, Neumeyer and Pilz (2006) used this concept to define (smooth) monotone estimates, while Chernozhukov, Fernández-Val and Galichon (2010), Dette and Volgushev (2008) successfully applied rearrangements techniques to define quantile regression estimates without crossing. Recently, Dette and Wu (2019) used monotone rearrangements to detect relevant changes in a (not necessarily monotone) trend of a non-stationary time series.

Our paper is organized as follows. In Section 2, we recall the concept of monotone rearrangements and introduce our transformation of a given dependence measure to a new measure with the desired properties (1.1)–(1.3) in several steps. First, we characterize the dependence measure \(\mu(X, Y) = \mu(C)\) in terms of the copula \(C\) of the corresponding distribution function of \((X, Y)\). Then we apply a monotone rearrangement to the partial derivative of \(C\) with respect to its first argument, which essentially constitutes the conditional distribution\(^1\) \(u \mapsto \mathbb{P}(F_Y(Y) \leq v \mid F_X(X) = u)\), and integrate it with respect to the conditioning coordinate. The resulting rearranged copula is denoted by \(C^\uparrow\) and, roughly speaking, it can be shown that the *rearranged dependence measure*

\[R_\mu(C) := \mu(C^\uparrow)\]

satisfies the desired properties (1.1)–(1.3). In Section 3, we propose an estimate of the rearranged dependence measure \(R_\mu(C)\), which is obtained by applying the procedure to the so-called checkerboard

\(^1\)\(F_X\) and \(F_Y\) denote the marginal distributions of \(X\) and \(Y\), respectively.
copula (see Li et al., 1997, for example). We also prove consistency of the estimate and illustrate the finite sample properties of our approach by means of a small simulation study in Section 4. Finally, all proofs are deferred to appendices and the online supplement which also contains some general results on monotone rearrangements, used for our theoretical arguments.

2. Dependence measures with properties (1.1)–(1.3)

In this section, we construct a rearranging transformation which assigns to some given dependence measure $\mu$ a new measure $R_\mu$ with the desired properties (1.1)–(1.3). We also discuss some further useful properties of the rearranged measure. To be precise, let $(X, Y)$ denote a 2-dimensional random vector with continuous distribution function $F$ and marginal distribution functions $F_X$ and $F_Y$. The dependence structure of $X$ and $Y$ is completely encoded in the (unique) copula $C = C_{X,Y}$ (see Definition S.I.1 in the Supplementary Material (Strothmann, Dette and Siburg (2022))) defined by the equation

$C(F_X(x), F_Y(y)) = F(x,y)
$ as described, for instance, in Nelsen (2006). The class of all copulas corresponding to continuous 2-dimensional distributions is denoted by $C$.

The proofs of all the results in this section are deferred to Appendix A and the supplementary material.

2.1. New dependence measures by monotone rearrangements

We always consider dependence measures of $(X, Y)$ as functions of the copula $C = C_{X,Y}$ and consequently use the notations $\mu(X, Y)$ and $\mu(C)$ interchangeably. The key ingredient is a rearrangement of the conditional distribution functions $u \mapsto P(F_Y(Y) \leq v \mid F_X(X) = u) = \partial_1 C(u, v) := \frac{\partial}{\partial u} C(u, v)$ (2.1) of the vector $(F_X(X), F_Y(Y))$. Note that the partial derivative $\partial_1 C(u, v)$ is only defined almost everywhere. We will suppress this fact in our notation for the remainder of this article.

**Definition 2.1.** A copula $C \in C$ is called stochastically increasing (resp. decreasing) if $u \mapsto \partial_1 C(u, v)$ is decreasing (resp. increasing) for each $v$. The class of all stochastically increasing copulas is denoted by $C^\uparrow$. A copula $C$ is called stochastically monotone if it is either stochastically increasing or decreasing. Similarly, a random variable $Y$ is stochastically increasing (resp. decreasing/monotone) in $X$ if $C_{XY}$ is stochastically increasing (resp. decreasing/monotone).

We will now introduce a procedure transforming an arbitrary copula into a stochastically increasing one. It is based on the monotone rearrangement of a univariate function, which is a classical concept in majorization theory (see, for example, Bennett and Sharpley, 1988, Chong and Rice, 1971). Namely, if $\lambda$ denotes the Lebesgue measure and $f : [0, 1] \to \mathbb{R}$ is a Borel measurable function, then the decreasing rearrangement $f^* : [0, 1] \to \mathbb{R}$ of $f$ is defined by

$f^*(t) := \inf\{x \mid \lambda(\{u \in [0, 1] \mid f(u) > x\}) \leq t\}.
$ (2.2)

Obviously, the function $f^*$ is a decreasing function and we have $f^* = f$ whenever $f$ is decreasing and right-continuous.
**Definition 2.2.** The *stochastically increasing rearrangement*, (SI)-rearrangement in short, of a copula \( C \in C \) is defined as

\[
C^\dagger(u, v) := \int_0^u (\partial_1 C)^*(s, v) \, ds
\]  

(2.3)

where the rearrangement (2.2) is applied to the first coordinate of \( \partial_1 C(u, v) \).

Our next result shows that \( C^\dagger \) defines in fact a copula.\(^2\)

**Theorem 2.3.** *The (SI)-rearrangement \( C^\dagger \) of a copula \( C \) is a stochastically increasing copula. Moreover, \( C^\dagger = C \) if and only if \( C \) is stochastically increasing itself.*

For a given dependence measure \( \mu \), we now define a new dependence measure by

\[
R\mu(C) := \mu(C^\dagger).
\]

(2.4)

We call \( R\mu \) the *rearranged dependence measure* obtained from \( \mu \). Note that, in general, \( R\mu \) differs from \( \mu \) and hence yields a new measure of dependence. Our main result is the following.

**Theorem 2.4.** Suppose \( \mu \) is a dependence measure which, when restricted to the set \( C^\dagger \), satisfies the properties (1.1)–(1.3). Then the rearranged dependence measure \( R\mu \) satisfies the properties (1.1)–(1.3) on the whole set \( C \).

**Remark 2.5.** Recently, dependence measures with the properties (1.1)–(1.3) have found considerable attention in the literature. For example, Trutschnig (2011) defined the measure

\[
\zeta_1(C) = 3 \int_0^1 \int_0^1 |\partial_1 C(u, v) - v| \, du \, dv,
\]

while Dette, Siburg and Stoimenov (2013) and Chatterjee (2021) considered (and proposed estimates for) the measure

\[
r(C) = 6 \int_0^1 \int_0^1 (\partial_1 C(u, v) - v)^2 \, du \, dv.
\]

(2.5)

It will be shown in Appendix A that the stochastically increasing rearrangement captures the entire information about the degree of dependence as defined by these measures in the sense that

\[
\zeta_1(C) = \zeta_1(C^\dagger) \quad \text{as well as} \quad r(C) = r(C^\dagger).
\]

(2.6)

**2.2. Examples**

In this section, we illustrate the rearrangement approach by a couple of examples. In particular, our method is applicable to construct a rearranged Spearman’s \( \rho \) or Kendall’s \( \tau \) from those classical measures of concordance. Moreover, we derive some interesting properties of the rearranged dependence measures.

\(^2\)The analogous definition of the stochastically decreasing rearrangement copula \( C^\ddagger \) is given and discussed in Appendix A; see also (Ansari and Rüschendorf, 2021). All subsequent theoretical results can be stated and proven for either \( C^\dagger \) or \( C^\ddagger \).
Example 2.6 (Schweizer-Wolff measures). Let $\Pi(u, v) = uv$ denote the independence copula. Each $L^p$-norm with $1 \leq p < \infty$ defines a so-called Schweizer-Wolff measure

$$
\sigma_p(C) := \frac{\|C - \Pi\|_p}{\|C^+ - \Pi\|_p},
$$

where the copula $C^+$ is defined by $C^+(u, v) = \min\{u, v\}$ (see S.I in the Supplementary Material (Strothmann, Dette and Siburg (2022))). The measure $\sigma_1$ was considered in Schweizer and Wolff (1981), $\sigma_2$ is also known as Blum-Kiefer-Rosenblatt’s $R$, and the general case $p \geq 1$ can be found in Section 5.3.1 of Nelsen (2006). It is easy to see that properties (1.1) and (1.2) hold for $\sigma_p$, and it is well known that $\sigma_p(C) = 1$ if and only $Y = f(X)$ for some strictly monotone (but not just measurable) function $f$ (Nelsen, 2006, Sect. 5.3.1). Consequently, $\sigma_p$ does not satisfy property (1.3). On the other hand, it will be shown in Appendix A that the properties (1.1)–(1.3) do hold for the restriction of $\sigma_p$ to the set $C^\uparrow$. Therefore, the rearranged Schweizer-Wolff measure

$$
R_{\sigma_p}(C) = \frac{\|C^\uparrow - \Pi\|_p}{\|C^\uparrow - \Pi\|_p}
$$

defines a new dependence measure on $C$ satisfying all the properties (1.1)–(1.3) on $C$.

Example 2.7 (Measures of concordance). Let $\kappa : C \to [-1, 1]$ be a measure of concordance (see Definition S.I.5 in the Supplementary Material (Strothmann, Dette and Siburg (2022))). Typical examples include Spearman’s $\rho$, Kendall’s $\tau$, Gini’s $\gamma$, and Blomqvist’s $\beta$ (see Appendix A for a representation of these measures in terms of the copula). We will prove in Appendix A that the measures $\rho, \tau$ and $\gamma$ satisfy (1.1)–(1.3) on the set $C^\uparrow$ (but not on $C$); on the other hand, Blomqvist’s $\beta$ does not satisfy (1.3) on $C^\uparrow$.

Consequently, by Theorem 2.4, the rearranged Spearman’s $\rho$ ($R_\rho$), Kendall’s $\tau$ ($R_\tau$) and Gini’s $\gamma$ ($R_\gamma$) define dependence measures (different from their original measures) satisfying (1.1)–(1.3) on $C$.

We will now see that, surprisingly, the Schweizer-Wolff measure $\sigma_1$ and Spearman’s $\rho$ induce the same rearranged dependence measure.

Proposition 2.8. We have $R_{\sigma_1} = R_\rho$.

Remark 2.9. We point out that there are even uncountably many dependence measures $\mu$ satisfying $R_\mu = R_\rho$. Indeed, pick any function $f : C \to [0, 1]$ being 1 on $C^\uparrow$ and 0 outside some neighbourhood of $C^\uparrow$ (apply Urysohn’s lemma to the closed convex set $C^\uparrow$), and consider the dependence measures

$$
\mu := f \rho + (1 - f) \nu
$$

where $\nu \neq \rho$. Then $\mu = \rho$ on $C^\uparrow$ so that $R_\mu = R_\rho$, regardless of the choice of $\nu$.

Remark 2.10. A referee raised the question if there exist “well-known” dependence measure $\mu \neq r$ such that $R_\mu = R_r$. In the following, we will derive a necessary condition for such measures. Since $r$ is invariant under rearrangement, we have $R_r(C) = r(C)$. Now suppose $X$ and $Y$ follow a bivariate normal distribution with correlation $p \in [-1, 1]$ and corresponding copula $C_p$. Since $C_p = C^\uparrow_p$, it follows that for a normal distribution the dependence measure $\mu$ must satisfy

$$
\mu(X, Y) = \mu(C_p) = \mu(C^\uparrow_p) = R_\mu(C_p) = R_r(C_p) = r(C_p) = \frac{3}{\pi} \arcsin \left( \frac{1 + p^2}{2} \right) - \frac{1}{2}.
$$
We are not aware of any “well-known” dependence measure fulfilling this property. However, by the same technique as in Remark 2.9, it can be shown that there exist infinitely many “dependence measures” $\mu$ on $C$ such that $R_{\mu} = R_r$. Thus, the equivalence class $[r] := \{\mu \mid R_{\mu} = R_r\}$ is not a singleton.

While a measure of concordance $\kappa$ measures the strength of the monotone association between two random variables, the corresponding rearranged dependence measure $R_{\kappa}$ measures the strength of their (directed) functional relationship. Thus $\kappa$ should always attain smaller values than $R_{\kappa}$. This heuristic is confirmed by the next theorem, which applies, in particular, to Spearman’s $\rho$ and Kendall’s $\tau$.

**Theorem 2.11.** Let $\kappa$ be a measure of concordance satisfying (1.1)–(1.3) on the set $C^\uparrow$. Then

$$|\kappa(C)| \leq R_{\kappa}(C)$$

for all $C \in C$, with equality whenever $C$ is stochastically monotone.

**Remark 2.12.** The inequality (2.8) connecting the underlying measure $\mu$ and $R_{\mu}$ can be extended beyond concordance measures. Whenever the measure $\mu$ is ordered with respect to the pointwise ordering of copulas and fulfills for all random variables $X$ and $Y$ either $\mu(1 - X, Y) = -\mu(X, Y)$ or $0 \leq \mu(X, Y)$, then $|\mu(X, Y)| \leq R_{\mu}(X, Y)$.

### 2.3. Data processing inequality and self-equitability

Informally, the so-called data processing inequality states that a (random or functional) modification of the input data cannot increase the information contained in the data; see, e.g., Cover and Thomas (2006) for an in-depth treatment of the data processing inequality in the context of information theory.

We assume in the following that the dependence measure $\mu$ is monotone on $C^\uparrow$ with respect to the pointwise order, i.e. we have

$$C_1 \leq C_2 \implies \mu(C_1) \leq \mu(C_2)$$

for all $C_1, C_2 \in C^\uparrow$. Note that this monotonocity condition holds for many dependence measures. For example, (2.9) is satisfied for any concordance measure (see Definition S.I.5 for a precise definition), the Schweizer-Wolff measures $\sigma_p$ in (2.7) as well as the measures of complete dependence $\zeta_1$ and $r$ introduced in Remark 2.5.

**Proposition 2.13 (Data processing inequality).** Assume that the dependence measure $\mu$ satisfies (2.9), and let $X, Y, Z$ be continuous random variables such that $Y$ and $Z$ are conditionally independent given $X$. Then the data processing inequality

$$R_{\mu}(Z, Y) \leq R_{\mu}(X, Y)$$

holds. In particular, $R_{\mu}(f(X), Y) \leq R_{\mu}(X, Y)$ holds for all\(^3\) measurable functions $f$.

Similar to (Geenens and Lafaye de Micheaux, 2022, Proposition 2.1), the data processing inequality also immediately yields an asymmetric version of the so-called self-equitability introduced in Kinney and Atwal (2014).

\(^3\)Note that for $R_{\mu}(f(X), Y)$ to be well-defined, $f(X)$ needs to be a continuous random variable.
Rearranged dependence measures

**Corollary 2.14.** Assume that $\mu$ satisfies (2.9). If $f$ is a measurable function such that $X$ and $Y$ are conditionally independent given $f(X)$, then

$$R_\mu(f(X), Y) = R_\mu(X, Y).$$

In particular, $R_\mu(g(X), Y) = R_\mu(X, Y)$ holds for all measurable bijections $g$.

Intuitively, Corollary 2.14 states that, in a regression model $Y = f(X) + \epsilon$, the dependence measure $R_\mu(X, Y)$ depends only on the strength of the noise $\epsilon$ and not on the specific form of $f$. A similar idea is illustrated in Figures 3 and 4 of Junker, Griessenberger and Trutschnig (2021).

### 2.4. Multivariate rearranged dependence measures

In this section we explain how the rearrangement technique can be generalized to a multivariate setting as follows. For any measure $m$ on $[0, 1]^d$ and any Borel measurable function $f : [0, 1]^d \to \mathbb{R}$, the *decreasing rearrangement* $f^*$ of $f$ is defined by

$$f^*(t) := \inf\{x \in \mathbb{R} \mid m(\{u \in [0, 1]^d \mid f(u) > x\}) \leq t\}.$$

As in the former case $d = 1$, $f^*$ is always a decreasing (univariate) function.

Now, let $(X, Y)$ denote a $(d + 1)$-dimensional random vector with continuous distribution function $F$ and marginal distribution functions $F_{X_1}, \ldots, F_{X_d}$ and $F_Y$. Using the disintegration approach introduced by Griessenberger, Junker and Trutschnig (2022), for each $(d + 1)$-copula $C$ there exists a Markov kernel $K_C : [0, 1]^d \times \mathcal{B}([0, 1]) \to [0, 1]$ such that

$$C(u, v) = \int_{[0,u]} K_C(s, [0,v]) \, d\mu_{C^{1\ldots d}}(s),$$

where $C^{1\ldots d}(u_1, \ldots, u_d) := C(u_1, \ldots, u_d, 1)$ denotes the marginal copula with its induced measure $\mu_{C^{1\ldots d}}$. Similar to Theorem 2.3,

$$C^\uparrow(u, v) := \int_0^d (K_C)^*(s, v) \, ds \quad (2.10)$$

is again a bivariate stochastically increasing copula, where the rearrangement is applied to the measure $m = \mu_{C^{1\ldots d}}$ and the function $s \mapsto K_C(s, [0,v])$ for every fixed $v \in [0, 1]$. Note that in the case $d = 1$ the representation (2.10) reduces to (2.3) since $C^\uparrow(u_1) := C(u_1, 1) = u_1$ and $K_C(s, [0,v]) = \partial_1 C(s, v)$ almost everywhere.

Given any bivariate dependence measure $\mu$, the rearranged dependence measure $R_\mu(C) := \mu(C^\uparrow)$ is now a multivariate measure of dependence in the sense that the multivariate versions

(M 1.1) $0 \leq R_\mu(C) \leq 1$
(M 1.2) $R_\mu(C) = 0$ if, and only if, $X$ and $Y$ are independent
(M 1.3) $R_\mu(C) = 1$ if, and only if, $Y = f(X)$ for some measurable function $f : \mathbb{R}^d \to \mathbb{R}$

of (1.1)–(1.3) hold for every $(d + 1)$-copula $C$. We finally note that the multivariate rearrangement and the induced rearranged dependence measures enjoy many properties known from other multivariate measures of complete dependence. For example, the multivariate rearrangement fulfils the information gain inequality

$$C^\uparrow_{X_{i_1}, Y} \leq C^\uparrow_{(X_{i_1}, X_{i_2}), Y} \leq \cdots \leq C^\uparrow_{(X_{i_1}, \ldots, X_{i_d}), Y},$$
and this also holds for $R_\mu(C)$ if the dependence measure $\mu$ is monotone with respect to the pointwise ordering of copulas. Moreover, if $X_2, \ldots, X_d$ and $Y$ are conditionally independent given $X_1$, we have $C_{(X_1, \ldots, X_d), Y}^\uparrow = C_{X_1, Y}^\uparrow$.

3. Approximation and estimation

In general, the computation of the rearrangement of a function, and hence the computation of $C^\uparrow$, may be a difficult task. In this section, we discuss techniques to approximate $C^\uparrow$ and $R_\mu(C)$ and to estimate the rearranged dependence measure $R_\mu$ from a sample of independent and identically distributed observations $(X_1, Y_1), \ldots, (X_n, Y_n)$. In principle, one would like to estimate the copula $C$ through a “smooth” statistic, say $\hat{C}_n$, and then apply Definition 2.2 to calculate the rearrangement $\hat{C}_n^\uparrow$ and the rearranged dependence measure $R_\mu(\hat{C}_n) = \mu(\hat{C}_n^\uparrow)$. \hfill (3.1)

While various smooth estimators have been proposed (see Chen and Huang, 2007, Fermanian, Radulović and Wegkamp, 2004, Genest, Nešlehová and Rémillard, 2017, Omelka, Gijbels and Veraverbeke, 2009, among others), the simultaneous estimation of the rearrangement poses various difficulties. We will now propose a simple solution to this problem.

Our approach is based on an approximation scheme for $C^\uparrow$ in the theoretical as well as empirical setting using the concept of checkerboard copulas, thereby circumventing the need to treat partial derivatives explicitly. Checkerboard copulas are an important tool in statistical applications; for a detailed discussion we refer, among others, to Genest, Nešlehová and Rémillard (2017) and Junker, Griessenberger and Trutschnig (2021). To be precise let $A = a_{k\ell}$ be a matrix with entries $a_{k\ell}$ satisfying

$$a_{k\ell} \geq 0 \quad \text{for all } k = 1, \ldots, N_1 \text{ and } \ell = 1, \ldots, N_2,$$

$$\sum_{k=1}^{N_1} a_{k\ell} = N_1 \quad \text{for all } \ell = 1, \ldots, N_2,$$

$$\sum_{\ell=1}^{N_2} a_{k\ell} = N_2 \quad \text{for all } k = 1, \ldots, N_1.$$ \hfill (3.2)

Then the function $C_{N_1, N_2}(A) : [0, 1]^2 \rightarrow [0, 1]$ defined by

$$C_{N_1, N_2}(A)(u, v) := \sum_{k,\ell=1}^{N_1,N_2} a_{k\ell} \int_0^u \mathbb{1}_{\left[\frac{k-1}{N_1}, \frac{k}{N_1}\right]}(s) \, ds \int_0^v \mathbb{1}_{\left[\frac{\ell-1}{N_2}, \frac{\ell}{N_2}\right]}(t) \, dt$$ \hfill (3.3)

is a copula and called the checkerboard copula of the matrix $A$. For a copula $C$ (see Definition S.1.1) its induced checkerboard copula is defined as

$$C_{N_1, N_2}(C) := C_{N_1, N_2}(A_{N_1, N_2}) ,$$ \hfill (3.4)

where the elements of the doubly stochastic matrix $A_{N_1, N_2}$ are given by

$$(A_{N_1, N_2})_{k\ell} := N_1 N_2 \cdot V_C \left( \left[ \frac{k-1}{N_1}, \frac{k}{N_1} \right] \times \left[ \frac{\ell-1}{N_2}, \frac{\ell}{N_2} \right] \right)$$ \hfill (3.5)
and $V_C(B)$ denotes the measure of the (Borel-)set $B \subset [0,1]^2$ induced by the copula $C$.

Note that in contrast to most of the literature, we define a (empirical) checkerboard copula also for non-square matrices $A$ satisfying (3.2). For $N = N_1 = N_2$ the representation (3.3) essentially reduces, up to a scaling factor $N$, to the common definition based on doubly stochastic square matrices (see Genest, Nešlehová and Rémillard, 2017, Junker, Griessenberger and Trutschnig, 2021). The consideration of the rectangular case, however, is necessary to address asymmetric dependencies between $X$ and $Y$ resp. $Y$ and $X$.

We point out that the partial derivatives of the copula $C_{N_1,N_2}^\#(A)$ in (3.3) are piecewise constant for fixed $v \in [0,1]$ with

$$
\partial_1 C_{N_1,N_2}^\#(A) \left( \frac{u_j}{N_2} \right) = \frac{1}{N_2} \sum_{k=1}^{j} a_{k\ell} \quad \text{for } u \in \left[ \frac{k-1}{N_1}, \frac{k}{N_1} \right].
$$

Thus, the (SI)-rearrangement satisfies $C_{N_1,N_2}^\#(A) = C_{N_1,N_2}^\#(A)$ if and only if

$$
\sum_{j=1}^{\ell} a_{k_2,j} \leq \sum_{j=1}^{\ell} a_{k_1,j}
$$

for all $1 \leq \ell \leq N_2$ and all $1 \leq k_1 \leq k_2 \leq N_1$. In other words, $C_{N_1,N_2}^\#(A) = C_{N_1,N_2}^\#(A)$ if and only if the rows of $A$ are ordered with respect to the majorization ordering of vectors (see Marshall and Arnold, 2011). This suggests the following Algorithm 1 for calculating the (SI)-rearrangement (as defined in Definition 2.2) of an arbitrary checkerboard copula.

**Algorithm 1:** Rearranged checkerboard copula

**Data:** matrix $A \in \mathbb{R}^{N_1 \times N_2}$ with entries satisfying (3.2)

**Result:** (SI)-rearrangement $C_{N_1,N_2}^\#(A)^\dagger$ of the checkerboard copula $C_{N_1,N_2}^\#(A)$

1. Calculate $B^\ell_k := \sum_{j=1}^{\ell} a_{kj}$ and set $B^0_k := 0$.
2. For every $\ell = 0, \ldots, N_2$, sort $B^\ell_k$ in a decreasing order and denote the result by $\tilde{B}^\ell_k$.
3. Calculate $a_{k\ell}^\dagger$ iteratively using

$$
a_{k\ell}^\dagger := \tilde{B}^\ell_k - \tilde{B}^{\ell-1}_k \geq 0.
$$

4. Define $A^\dagger := (a_{k\ell}^\dagger)_{k=1,\ldots,N_1}$ and

$$
C_{N_1,N_2}^\#(A)^\dagger := C_{N_1,N_2}^\#(A^\dagger).
$$

**Theorem 3.1.** For any matrix $A \in \mathbb{R}^{N_1 \times N_2}$ satisfying (3.2), the function $C_{N_1,N_2}^\#(A)^\dagger$ defined in Algorithm 1 is the (SI)-rearrangement of the checkerboard copula $C_{N_1,N_2}^\#(A)$.

We now turn to the estimation of the population dependence measure $R_\mu(C) = \mu(C^\dagger)$ from a sample of independent and identically distributed observations. Because there exists in general no analytic expression for $R_\mu(C)$, this is a challenging task and we proceed in two steps. First, note that the population measure $R_\mu(C)$ can be approximated by $R_\mu(C_{N_1,N_2}^\#(C))$ using the induced checkerboard copula
$C_{N_1,N_2}^\#(C)$ of $C$ defined in (3.4) since

$$C_{N_1,N_2}^\#(C) \uparrow C \uparrow$$

(3.7)

where $C_{N_1,N_2}^\#(C) \uparrow$ denotes the rearrangement of $C_{N_1,N_2}^\#(C)$. Secondly, we replace the unknown weights in (3.5) by corresponding estimates to obtain an empirical checkerboard copula, which is then rearranged by Algorithm 1.

We begin with the approximation of $C \uparrow$ by the rearranged induced checkerboard copula. Since it is well known that the pointwise convergence is unable to capture complete dependence (see Mikusiński, Sherwood and Taylor, 1992), we consider the finer metrics

$$D_p(C_1,C_2) := \left( \int_0^1 \int_0^1 |\partial_1 C_1(u,v) - \partial_1 C_2(u,v)|^p \, du \, dv \right)^{\frac{1}{p}}$$

for $1 \leq p < \infty$ introduced in Trutschnig (2011).

**Theorem 3.2.** For any copula $C$, the rearranged induced checkerboard copula $C_{N_1,N_2}^\#(C) \uparrow$ converges to the rearranged copula $C \uparrow$ with respect to $D_p$, i.e.

$$D_p(C_{N_1,N_2}^\#(C) \uparrow, C \uparrow) \to 0$$

as $N_1,N_2 \to \infty$. In particular, $C_{N_1,N_2}^\#(C) \uparrow$ converges uniformly towards $C \uparrow$.

In order to carry over the convergence of $C_n \uparrow$ to $C \uparrow$ and establish consistency of the estimator, we require that the underlying dependence measure $\mu$ is continuous on $C \uparrow$ with respect to pointwise convergence, i.e. that

$$C_n \to C \implies \mu(C_n) \to \mu(C)$$

(3.8)

holds for all copulas $C_n, C \in C \uparrow$. We point out that most classical measures are continuous in this sense. In fact, any concordance measure (see Definition S.I.5), the Schweizer-Wolff measures $\sigma_p$ in (2.7), as well as the measures of complete dependence $\zeta_1$ and $r$ in Remark 2.5 fulfil our continuity condition$^4$.

**Theorem 3.3.** If the dependence measure $\mu$ satisfies (3.8) then

$$R_{\mu}(C_{N_1,N_2}^\#(C)) \to R_{\mu}(C) \quad \text{as } N_1,N_2 \to \infty.$$ 

Next, we consider a random sample of independent identically distributed observations $(X_1,Y_1), \ldots, (X_n,Y_n)$. Similar to Li, Mikusiński and Taylor (1998) and Junker, Griessenberger and Trutschnig (2021), who considered the case $N_1 = N_2$, we define the empirical checkerboard copula with bandwidth $N_1,N_2 < n$ by

$$C_{N_1,N_2}^\# := C_{N_1,N_2}^\#(C_{n,n}(\hat{A}_n)),$$  

(3.9)

where $\hat{A}_n = (\hat{a}_{ij})$ is the $n \times n$ permutation matrix defined by

$$\hat{a}_{ij} := \begin{cases} 1 & \text{if there exists some } k \text{ with rank}(X_k) = i \text{ and rank}(Y_k) = j \\
0 & \text{else} \end{cases}$$

$^4$For $\zeta_1$ and $r$ this follows from (Siburg and Strothmann, 2021, Prop. 3.6).
and \( \text{rank}(x_k) \) denotes the rank of \( x_k \) among \( x_1, x_2, \ldots, x_n \). Finally, we define

\[
\hat{R}_\mu := R_\mu(\hat{C}^\#_{N_1, N_2, n})
\]

(3.10)
as an estimator of \( R_\mu(C) \), which will be called rearranged \( \mu \)-estimate throughout this paper. The following result shows strong consistency of \( \hat{R}_\mu \).

**Theorem 3.4.** Assume that the dependence measure \( \mu \) fulfils the assumption (3.8), and let \((X_1, Y_1), \ldots, (X_n, Y_n)\) denote independent identically distributed random variables with a continuous distribution. If \( N_1 := [n^{s_1}] \), \( N_2 := [n^{s_2}] \) with \( s_1, s_2 \in (0, 1/2) \), then the estimator defined by (3.10) satisfies

\[
\hat{R}_\mu \to R_\mu(C) \text{ a.s. as } n \to \infty.
\]

**Remark 3.5.** For big data applications the time complexity of the new estimators is of importance. The calculation of the ranks requires \( O(n \log(n)) \) operations, while, in the absence of ties, the empirical checkerboard copula can be obtained by \( O(n) \) operations. The rearrangement in Algorithm 1 can be done by

\[
O(N_1 N_2) + O(N_2 N_1 \log(N_1)) + O(N_1 N_2) = O(n^{s_1} n^{s_2} \log(n^{s_1})) = O(n^{s_1 + s_2} \log(n))
\]
operations, where \( s_1 + s_2 < 1 \). Naturally, the time complexity to compute the dependence measure depends on the underlying measure but oftentimes requires \( O(N_1 N_2) = O(n^{s_1 + s_2}) \) operations. As a result, for fixed \( s_1, s_2 \) the time complexity is of order \( O(n \log(n)) \), which coincides with the complexity of Chatterjee’s estimator.

### 4. Finite sample properties

For a good performance of the estimate \( \hat{C}^\#_{N_1, N_2, n} \), an appropriate choice of the bandwidths \( N_1, N_2 \) will be crucial. These tuning parameters depend sensitively on the form of the underlying unknown copula, and for the finite sample illustrations presented below, we will use the following cross validation principle, which is adapted from density estimation (see, for example Stone, 1984).

Recall the definition of the empirical checkerboard copula \( \hat{C}^\#_{N_1, N_2, n} \), and denote its corresponding density by

\[
\hat{C}_{N_1, N_2, n}(u, v) := \frac{\partial^2}{\partial u \partial v} \hat{C}^\#_{N_1, N_2, n}(u, v).
\]

(3.11)

We define

\[
\text{CV}(N_1, N_2, n) := \int_0^1 \int_0^1 \hat{c}^2_{N_1, N_2, n}(u, v) \, du \, dv - \frac{2}{n} \sum_{i=1}^n \hat{c}^{-\mu}_{N_1, N_2, n-1}(\hat{U}_i, \hat{V}_i),
\]

where \( \hat{c}^{-\mu}_{N_1, N_2, n-1} \) denotes the estimator in (3.11) calculated from the data

\((X_1, Y_1), \ldots, (X_{i-1}, Y_{i-1}), (X_{i+1}, Y_{i+1}), \ldots, (X_n, Y_n)\)

and \( \hat{U}_i = \frac{1}{n+1} \sum_{j=1}^n I(X_j \leq X_i) \) and \( \hat{V}_i = \frac{1}{n+1} \sum_{j=1}^n I(Y_j \leq Y_i) \) are the normalized ranks of \( X_i \) and \( Y_i \) among \( X_1, \ldots, X_n \) and \( Y_1, \ldots, Y_n \), respectively. The data adaptive choice of the parameters \( N_1 \) and \( N_2 \) is defined as the minimizer of \( \text{CV}(N_1, N_2, n) \) with respect to \( N_1, N_2 \in \{[n^{1/4}], \ldots, [n^{1/2}]\} \). Note that, although consistency of the empirical checkerboard copula holds for \([n^{s_i}]\) with \( s_i \in (0, 1/2) \) (see Junker,
Griessenberger and Trutschnig (2021), we perform cross-validation only for $s_1, s_2 \in (1/4, 1/2)$. On the one hand, this saves computational time. On the other hand, for $s < 1/4$, the number of grid divisions is extremely small yielding almost no discernible influence on the outcome of the cross-validation procedure. Moreover, in cases, where the set of possible bandwidths is very large, we calculate the minimizer on the set $\{[n^{1/4}], [n^{1/4}] + 2, \ldots, [n^{1/2}]\}$ in order to save further computational time.

4.1. Simulation study

In this section, we present results from a simulation study investigating the performance of the estimator $\hat{R}_\mu$ defined in (3.10). All simulations have been conducted using the statistical software “R” (see R Core Team, 2021) and are based on 1000 replications in each scenario. The code used in the simulation study can be found at https://github.com/ChristopherStrothmann/RDM. Therein, the package “qad” (see Kasper et al., 2022) was used in a slightly adapted form to calculate the matrix $\hat{A}_n$, which is required for the definition of the empirical checkerboard copula in (3.9). As sample sizes we considered $n = 50, 100, 500$ and $1000$ and $N_1, N_2$ were chosen by the cross validation procedure described at the beginning of this section.

4.1.1. Stochastically increasing distributions

We begin with a study of the properties of the estimator (3.10) in the rather special case where the underlying copula is stochastically increasing. The corresponding samples have been generated using the package “copula” (see Hofert et al., 2020). As for stochastically monotone copulas we have $R_\mu = \mu$, we can calculate the dependence measure explicitly, and it is also reasonable to compare the new estimator $\hat{R}_\mu$ with commonly used estimators of $\mu$. The R-packages “XICOR” was used to estimate Chatterjee’s coefficient $\hat{\xi}$ of $r$ (see Chatterjee (2021)) and “qad” (see Kasper et al. (2022)) was used to estimate $\hat{\zeta}_1$ of Trutschnig’s $\zeta_1$.

The first two scenarios correspond to a 2-dimensional (centred) normal distribution with correlation matrix

$$\begin{pmatrix} 1 & p \\ p & 1 \end{pmatrix}$$

(3.12)

where $p = 0.25$ and $p = 0.75$, respectively. Since for $p > 0$, the corresponding copula, say $C_p$, is stochastically increasing, the rearranged Spearman’s $\rho$ equals

$$R_\rho(C_p) = R(X,Y) = \rho(X,Y) = \frac{6}{\pi} \arcsin \left(\frac{p}{2}\right) \quad (p \geq 0),$$

while the rearranged Kendall’s $\tau$ equals

$$R_\tau(C_p) = R_\tau(X,Y) = \tau(X,Y) = \frac{2}{\pi} \arcsin (p) \quad (p \geq 0).$$

The third example of a stochastically increasing copulas is a member of both the Archimedean and extreme-value copula families, which are widely applied, both theoretically as well as empirically. More precisely, we consider a Gumbel copula defined by

$$C_G^G(u,v) := \exp \left( - \left( (-\log u)^\theta + (-\log v)\theta \right)^{1/\theta} \right).$$
Rearranged dependence measures

Figure 1: Scatter plots of data (sample size \( n = 500 \)) from the Gaussian copula with correlation \( r = 0.25 \) (left panel), \( r = 0.75 \) (middle panel) and the Gumbel copula with parameter \( \theta = 3 \) (right panel).

where \( \theta > 1 \) denotes a parameter. Since the Gumbel copula is an extreme-value copula, it is stochastically increasing, where the rearranged Spearman’s \( \rho \) and Kendall’s \( \tau \) are given by

\[
R_\rho(C_\theta^G) = \rho(X,Y) = 12 \int_0^1 \frac{1}{(1 + (t^\theta + (1-t)^\theta)^{1/\theta})^2} dt - 3 \quad \text{and} \quad \tau(C_\theta^G) = \tau(X,Y) = \frac{\theta - 1}{\theta}.
\]

In Figure 1, we show scatter plots of data generated from the two Gaussian copulas \((p = 0.25, p = 0.75)\) and the Gumbel copula \((\theta = 3)\), where the sample size is \( n = 500 \). In the upper part of Table 1, we present the simulated mean and standard deviation of the rearranged estimate \( \hat{R}_\rho \), where \( \mu \) is either Spearman’s \( \rho \) (left part) or Kendall’s \( \tau \) (right part). Due to \( R_\mu(C) = \mu(C) \) for the three scenarios, the commonly used Spearman’s rank correlation coefficient \( \hat{\rho} \) and Kendall’s rank correlation coefficient \( \hat{\tau} \) can also be used to estimate \( R_\rho(C) \) and \( R_\tau(C) \), respectively. The corresponding results for these estimates are displayed in Table 1 as well (of course, in practice it is not known if the underlying copula is stochastically increasing).

We observe a reasonable behaviour of all rearranged estimates, which improves with increasing sample size. In general, there are only minor differences between the rearranged estimates \( \hat{R}_\rho \), \( \hat{R}_\tau \) and the non-rearranged estimates \( \hat{\rho} \), \( \hat{\tau} \), which are mainly caused by a slightly smaller bias of the non-rearranged estimates. For the Gaussian copula with correlation 0.25, the rearranged estimates \( \hat{R}_\rho \) and \( \hat{R}_\tau \) slightly overestimate their population version \( R_\rho \) and \( R_\tau \) if the sample size is \( n = 50 \) or 100. For all other scenarios, we observe an underestimation. The lower part of Table 1 shows some results for the complete dependence measures \( r \) and \( \zeta_1 \). We observe that the estimator \( \hat{R}_\zeta \) and Chatterjee’s estimator \( \xi \) behave very similar. On the other the estimator \( \hat{R}_\xi \) seems to have a smaller bias than \( \xi \) but at the cost of a larger variance.

4.1.2. A family of non-stochastically monotone distributions

In this section, we consider the more common situation where \( R_\mu \neq \mu \). To generate data from a family of 2-dimensional distributions with different degrees of dependence, let \( X \sim U(0,1) \) denote a uniformly (on the interval \([0,1]\)) distributed random variable and \( Z \sim N(0,1) \) a standard normal distributed random variable such that \( X \) and \( Z \) are independent. We consider the regression model

\[
Y := (X - 1/2)^2 + \sigma Z , \quad (3.13)
\]
Table 1. Simulated mean and standard deviation of various estimators of dependence measures. Upper part: rearranged Spearman’s $\rho$ estimate $\hat{R}_\rho$, (absolute) Spearman’s rank correlation coefficient $|\hat{\rho}|$ (left part), rearranged Kendall’s $\tau$ estimate $\hat{R}_\tau$ and (absolute) Kendall’s rank correlation coefficient $|\hat{\tau}|$ (right part). Lower part: Siburg-Dette-Stoimenov $r$ (i.e. Chatterjee’s $\xi$) estimate $\hat{R}_r$, Chatterjee’s estimator $\hat{\xi}$ (left part), Trutschnig $\zeta_1$ estimate $\hat{R}_{\zeta_1}$ and $\hat{\zeta}_1$ (right part). The distribution of $(X,Y)$ is given by a centred normal with correlation matrix (3.12) with copula $C_\rho$ and by a Gumbel copula $C_3^G$.

| copula | $n$ | $R_\rho$ | $|\hat{\rho}|$ | $R_\tau$ | $|\hat{\tau}|$ |
|--------|-----|---------|-------------|---------|-------------|
| $C_{0.25}$ | 50 | 0.276 (0.132) | 0.246 (0.129) | 0.161 | 0.185 (0.090) |
| | 100 | 0.263 (0.093) | 0.236 (0.096) | 0.176 (0.063) |
| | 500 | 0.224 (0.043) | 0.240 (0.043) | 0.150 (0.029) |
| | 1000 | 0.226 (0.030) | 0.239 (0.029) | 0.151 (0.020) |
| $C_{0.75}$ | 50 | 0.734 | 0.721 (0.075) | 0.540 | 0.473 (0.079) |
| | 100 | 0.694 (0.063) | 0.727 (0.051) | 0.496 (0.055) |
| | 500 | 0.714 (0.025) | 0.732 (0.023) | 0.517 (0.023) |
| | 1000 | 0.723 (0.017) | 0.734 (0.015) | 0.527 (0.015) |
| $C_3^G$ | 50 | 0.803 (0.057) | 0.839 (0.050) | 0.667 | 0.599 (0.058) |
| | 100 | 0.826 (0.040) | 0.844 (0.037) | 0.628 (0.044) |
| | 500 | 0.844 (0.016) | 0.848 (0.015) | 0.653 (0.019) |
| | 1000 | 0.847 (0.011) | 0.848 (0.010) | 0.659 (0.012) |

where $\sigma$ is a non-negative constant. Note that the correlation between $X$ and $Y$ is 0, by construction and that a similar model has been studied in Chatterjee (2021). Model (3.13) contains perfect functional dependence of $X$ and $Y$ (for $\sigma = 0$) and independence in the limit for $\sigma \rightarrow \infty$. The corresponding scatter plots from $n = 500$ independent observations according to model (3.13) with $\sigma = 0, 0.1$ and 0.3 are displayed in Figure 2, while the upper part of Table 2 shows the simulated mean and standard deviation of the estimates $\hat{R}_\rho$ (for the rearranged Spearman’s $\rho$) and $\hat{R}_\tau$ (for the rearranged Kendall’s $\tau$). For $\sigma > 0$ the “true” values of $R_\rho$ and $R_\tau$ have been obtained by simulation using a sample of size $n = 1000000$ and bandwidths $N_1 = N_2 = \lfloor n^{0.45} \rfloor$. The empirical results confirm the consistency statement in Theorem 3.4. In the table, we also display the simulated mean of the non-rearranged estimators $|\hat{\rho}|$ and $|\hat{\tau}|$, which do not yield reasonable results. In the lower part of Table 2 we show again some results for the complete dependence measures $r$ and $\zeta_1$. For $\sigma = 0.1$ and $\sigma = 0.3$ the differences
between the estimators $\hat{R}_r$ and $\hat{\xi}$ are again very small. However for complete dependence ($\sigma = 0$) Chatterjee’s estimator yields a better performance. In this case the estimators $\hat{R}_{\xi_1}$ and $\hat{\xi}_1$ exhibit a similar behaviour, while with increasing $\sigma$ the estimator $\hat{\xi}_1$ has a larger bias than $\hat{R}_{\xi_1}$ (but a slightly smaller variance).

### 4.2. Power analysis

In this section we compare different dependence measures when they are applied for independence testing, that is

$$H_0: X \text{ and } Y \text{ are independent}$$

For modeling the dependence structure we consider a Gaussian copula $C_p$ with $p \in [0, 1]$ and the copula corresponding to the non-stochastically monotone distribution in model (3.13) with parameter $1/\sigma$. All results are based on a sample of size $n = 200$ and 2000 simulation runs are used to calculate the rejection probabilities.

We consider tests based on estimators for the rearranged Spearman’s $\rho$ and rearranged Kendall’s $\tau$. As benchmark we also study the tests based on $\xi_1$ (see Junker, Griessenberger and Trutschnig (2021), Trutschnig (2011)) and $r$ (see Chatterjee (2021), Dette, Siburg and Stoimenov (2013)). For the tests based on $R_\rho$, $R_\tau$ and $\xi_1$ we use a permutation test with 1000 permutations. For the test based on $r$ we use Chatterjee’s test with quantiles from the asymptotic distribution. The simulated rejection probabilities are displayed in Figure 3 for the Gaussian copula (left panel) and model (3.13) (right panel). Note that the case of independence corresponds to the choice $p = 0$ and $1/\sigma \to 0$, while complete dependence is obtained for $p = 1$ and $1/\sigma = \infty$. In this case, the nominal level is well approximated by all 4 tests. With increasing $p$ and increasing $1/\sigma$ we model more dependence and we can study the power of the different tests. We observe no differences in the power of the tests based on $R_\rho$, $R_\tau$ and $\xi_1$. However, all tests outperform Chatterjee’s test.

### 4.3. Data example

In this section we briefly revisit a data example which was investigated by Chatterjee (2021) to study the performance of his correlation coefficient in the analysis of yeast gene expression data. The data consists of the expressions of 6223 yeast genes and was originally analyzed by Spellman et al. (1998)
who tried to identify genes whose transcript levels oscillate during the cell cycle. For each gene, the
gene expression was observed at 23 time points. Because the number of genes is large, visual inspection
is not possible and Reshef et al. (2011) proposed to use the MIC and MINE correlation coefficient to
analyze the data. Chatterjee (2021) compared the performance of his correlation coefficient with these
measures and demonstrated some advantages of his approach. We will now provide a brief illustration
analyzing this type of data with a rearranged dependence measure to demonstrate the ability of our
approach to also detect non-monotone dependencies. We begin with an analysis of the rearranged
Spearman’s rank coefficient $\hat{\rho}$. After that, we provide a very brief comparison of $\hat{\rho}$ with Chatterjee’s
correlation coefficient.

To be precise, we consider the curated data set (available through the R-package “minerva”) of 4381
genes. For each gene, we perform a permutation test based on Spearman’s rank correlation for the

**Table 2.** Simulated mean and standard deviation of various estimators of dependence measures. Upper part: rearranged Spearman’s $\rho$ estimate $\hat{\rho}$, (absolute) Spearman’s rank correlation coefficient $|\hat{\rho}|$ (left part), rearranged Kendall’s $\tau$ estimate $\hat{\tau}$, (absolute) Kendall’s rank correlation coefficient $|\hat{\tau}|$ (right part). Lower part: Siburg-Dette-Stoimenov $r$ (i.e. Chatterjee’s $\xi$) estimate $\hat{r}$, Chatterjee’s estimator $|\hat{r}|$ (left part), Trutschnig $\zeta_1$ estimate $\hat{r}_1$ and $|\hat{r}_1|$ (right part). The distribution of $(X,Y)$ is given by model (3.13).
hypotheses

\[ H_0 : \rho = 0 \quad \text{versus} \quad H_1 : \rho > 0 \]

and a permutation test based on the statistic \( \hat{R}_\rho \) for the hypotheses

\[ H_0 : R_\rho = 0 \quad \text{versus} \quad H_1 : R_\rho > 0, \tag{3.14} \]

where we use 10000 permutations. The corresponding \( p \)-values are use to identify the significant genes using the Benjamini–Hochberg FDR procedure with a false discovery rate of 0.05 (see, Benjamini and Hochberg, 1995). To concentrate on non-monotone dependencies, we exclude from those genes selected by the FDR procedure based on the rearranged Spearman’s rank correlation all genes which are also detected by Spearman’s rank correlation. This results in 84 remaining genes. In Figure 4 we display the transcript levels of the top 6 genes with the smallest \( p \)-values from the remaining data. We observe that the FDR procedure based on the rearranged Spearman’s rank correlation identifies additional dependencies, which are oscillating and are not found if the analysis is based on Spearman’s rank correlation. A similar observation was made by Chatterjee (2021) for his rank correlation coefficient, who used 4 alternative tests to exclude genes with a monotone behaviour (a gene was excluded, whenever one of these tests identified it as significant). Because both procedures are based on different dependence measures the finally identified 6 top genes do not necessarily coincide (only the gene YBL003C was selected by our and Chatterjee’s procedure). However, all 6 top genes found by Chatterjee (2021) are also selected by the FDR procedure based on rearranged Spearman’s rank correlation and vice versa. Moreover, the qualitative conclusion from both methods is same. Both methods are able to identify non-monotone (in the concrete example oscillating) associations.

We conclude with a brief comparison of the FDR procedures based on the rearranged Spearman’s and Chatterjee’s rank correlation coefficient, if they are used without sorting out monotone dependencies by preliminary analysis. In Figures 5 and 6, we display the transcript levels of the 6 genes with the smallest \( p \)-values after running the FDR procedure based on the two dependency measures. We
Figure 4: Transcript levels of the top genes, which were selected by the FDR procedure based on the rearranged Spearman's rank correlation coefficient, but not by the FDR procedure based on Spearman's rank correlation coefficient. (\( \alpha = 0.05 \)). The dashed lines represent the 3-nearest neighbour regression estimates.

observe again that both methods are able to identify non-monotone associations. Interestingly the top three genes identified by the rearranged Spearman’s rank correlation with the smallest three \( p \)-values exhibit an oscillating transcript level while it looks more monotone for the next three genes. For the FDR procedure based on Chatterjee’s rank correlation the picture is not so clear.

5. Conclusions and outlook

In this paper we developed a general strategy to transform a given dependence measure into a new one which exactly characterizes independence (with the value 0) as well as functional dependence (with the value 1). The approach is applicable to many of the commonly used dependence measures and we have also developed consistent estimates of the new dependence measures. An interesting question of future research is the asymptotic distribution of the new estimators. However, such an investigation will be very challenging because - in contrast to most of the literature - we do not consider “one specific”
Rearranged dependence measures

Figure 5: Transcript levels of genes, which were selected by the rearranged Spearman’s rank correlation coefficient. The figure shows the 6 top genes with the smallest p-values. The dashed lines represent the 3-nearest neighbour regression estimates.

estimator for “one specific” dependence measure, for which the asymptotic distribution is established. Thus one would have to identify classes of copulas and classes dependence measures for which such an asymptotic analysis is possible.

A further challenging question are rates of convergence of the new estimators. These rates are case-specific and will depend in an intricate way on the quality of the approximation of the copula $C$ by the induced checkerboard copula $C_{N,N,n}(C)$ (which might be different for different copulas) and on the dependence measure under consideration. To indicate how these rates can be obtained in principle, we consider exemplary the case of Spearman’s $\rho$ and note that

$$|R_\rho(C) - R_\rho(D)| = 12 \left| \int_{[0,1]^2} C^\uparrow(u,v) - D^\uparrow(u,v) \, d\lambda(u,v) \right|$$

$$= 12 \left| \int_{[0,1]^2} \int_0^u \partial_1 C^\uparrow(s,v) - \partial_1 D^\uparrow(s,v) \, ds \, d\lambda(u,v) \right|$$
Figure 6: Transcript levels of genes, which were selected by the Chatterjee’s correlation coefficient. The figure shows the 6 genes with the smallest \( p \)-values. The dashed lines represent the 3-nearest neighbour regression estimates.

\[
\begin{align*}
&\leq 12 \int_{[0,1]^2} \int_0^u \left| \partial_1 C^\uparrow(s, v) - \partial_1 D^\uparrow(s, v) \right| \, ds \, d\lambda(u, v) \\
&\leq 12 D_1(C^\uparrow, D^\uparrow) \leq 12 D_1(C, D) .
\end{align*}
\]

Thus, in order to obtain the rate for the rearranged estimate of Spearman’s \( \rho \) we need an estimate for \( D_1(C^\#_{N,N,n}, C) \). For this purpose we can use results of Junker, Griessenberger and Trutschnig (2021), who showed the inequality

\[
D_1(C^\#_{N,N,n}, C) \leq K \frac{\sqrt{\log \log n}}{n_1^{1/2-s}} + D_1(C^\#_{N,N,n}(C), C)
\]

(almost surely), and it remains to estimate the deterministic quantity \( D_1(C^\#_{N,N,n}(C), C) \). Estimates of this type are again case specific. For example, if \( C = \Pi \) is the independence copula it follows that
Almost surely for $0 < s < \frac{1}{2}$. Note that the distance $D_1$ appears here, because of the representation

$$\rho(C) = 12 \int_{[0,1]^2} C(u, v) \, dI(u, v) - 3$$

(see, e.g., Chapter 5 in Nelsen (2006)). A similar argument can be used to derive a rate for the estimator of the rearranged Blum-Kiefer-Rosenblatt’s $R$. Other dependence measures such as Kendall’s $\tau$ or Gini’s $\gamma$ have more complicated representations in terms of the copula $C$, which might require other metrics to bound $|R_\rho(C) - R_\rho(D)|$.

**Appendix A: Proofs of the results in Section 2**

**Proof of Theorem 2.3** In order to show that the stochastically increasing rearrangement, $C^\uparrow$ is a copula, we verify the properties (1) to (3) of Definition S.I.1.

1. It follows from $(\partial_1 C)^* (u, 0) = 0^* = 0$ that $C^\uparrow(u, 0) = 0$. The identity $C^\uparrow(0, v) = 0$ is trivial by Definition 2.2.
2. By definition, we have

$$C^\uparrow(u, 1) = \int_0^u (\partial_1 C)^*(s, 1) \, ds = \int_0^u 1^* \, ds = u.$$ 

In view of Proposition S.I.6(3), we further obtain that

$$C^\uparrow(1, v) = \int_0^1 (\partial_1 C)^*(s, v) \, ds = \int_0^1 (\partial_1 C)^*(\sigma_v(s), v) \, ds = \int_0^1 \partial_1 C(t, v) \, dt = v.$$ 

3. From Definition S.I.1(3) we see that $0 \leq \partial_1 C(\cdot, v_1) \leq \partial_1 C(\cdot, v_2)$ whenever $v_1 \leq v_2$. Combining this with Proposition S.I.6(2) yields $(\partial_1 C)^*(\cdot, v_1) \leq (\partial_1 C)^*(\cdot, v_2)$. Thus, the $C^\uparrow$-volume of a rectangle $[u_1, u_2] \times [v_1, v_2]$ satisfies

$$V_{C^\uparrow} ([u_1, u_2] \times [v_1, v_2]) = C^\uparrow(u_2, v_2) - C^\uparrow(u_1, v_2) - C^\uparrow(u_2, v_1) + C^\uparrow(u_1, v_1)$$

$$= \int_{u_1}^{u_2} (\partial_1 C)^*(s, v_2) - (\partial_1 C)^*(s, v_1) \, ds \geq 0.$$ 

Finally, we show that $C$ is stochastically increasing if and only if $C = C^\uparrow$. If $C = C^\uparrow$, of course, $C$ is stochastically increasing because $C^\uparrow$ is. Conversely, suppose $C$ is stochastically increasing, i.e., each $u \mapsto C(u, v)$ is concave. Then the right-hand derivative $u \mapsto \partial_1^+ C(u, v)$ is a decreasing and right-continuous function, and (Chong and Rice, 1971, Thm. 4.2) guarantees that $\partial_1^+ C(u, v) = (\partial_1 C)^*(u, v)$. This implies

$$C(u, v) = \int_0^u \partial_1^+ C(t, v) \, dt = \int_0^u (\partial_1 C)^*(t, v) \, dt = C^\uparrow(u, v).$$

**Proof of Theorem 2.4** We will require a preliminary result. For this, we first note that the so-called (SD)-rearrangement of $C$ defined by

$$C^\downarrow(u, v) := \int_0^u (\partial_1 C)^*(1 - s, v) \, ds = v - C^\uparrow(1 - u, v) = (C^- \ast C^\downarrow)(u, v)$$
is a stochastically decreasing copula.

**Lemma A.1.** For any copula $C$, we have $C^\downarrow(u, v) \leq C(u, v) \leq C^\uparrow(u, v)$.

**Proof.** By Theorem S.1.8(a) we obtain the upper estimate

$$C(u, v) = \int_0^1 \mathbb{I}_{[0, u]}(t) \partial_1 C(t, v) \, dt \leq \int_0^1 \mathbb{I}_{[0, u]}(t) (\partial_1 C)^\uparrow(t, v) \, dt = C^\uparrow(u, v).$$

The lower estimate follows analogously.

We will now prove properties (1.1)–(1.3) for $R_\mu(C) = \mu(C^\uparrow)$. For this, we say that the copula $C$ is *completely dependent* if there exists a measurable function $f$ such that $V = f(U)$. It is proven in Darsow, Nguyen and Olsen (1992) that $C$ is completely dependent if, and only if,

$$\partial_1 C(u, v) \in \{0, 1\}$$

(3.15)

for almost all $u \in [0, 1]$ and all $v \in [0, 1]$.

(1.1) Since $\mu$ only takes values between 0 and 1, we obtain the first assertion.

(1.2) If $C = \Pi$, we have $\mu(C^\uparrow) = \mu(\Pi) = 0$. If, on the other hand, $\mu(C^\uparrow) = 0$, we conclude $C^\uparrow = \Pi$ by the properties of $\mu$. But then $C^\uparrow = C^\uparrow \ast \Pi = \Pi$, and Lemma A.1 yields $\Pi = C^\uparrow \leq C \leq C^\uparrow = \Pi$, hence $C = \Pi$.

(1.3) If $C$ is completely dependent, then $C^\uparrow = C^\uparrow$ and $\mu(C^\uparrow) = \mu(C^\uparrow) = 1$ by definition. On the other hand, $\mu(C^\uparrow) = 1$ implies $C^\uparrow = C^\uparrow$ by the properties of $\mu$. Thus, $\partial_1 C(u, v) = (\partial_1 C)^\uparrow(\sigma, (u, v) \in \{0, 1\}$, so $C$ is completely dependent by (3.15).

**Proof of Equation (2.6)** The statement is an immediate consequence of the fact that the decreasing rearrangement of $g_\sigma(u) := \partial_1 C(u, v) - v$ is $g_\sigma^\uparrow(u) = \partial_1 C^\uparrow(u, v) - v$. As the decreasing rearrangement leaves all $L^p$-norms invariant, we conclude

$$\int_0^1 |\partial_1 C(u, v) - v|^p \, du = \int_0^1 |\partial_1 C^\uparrow(u, v) - v|^p \, du.$$

Integrating with respect to $v$ yields the desired result with $p = 1, 2$.

**Proof of the statements in Example 2.6** In this section we show that the Schweizer-Wolff measure $\sigma_p$ in (2.7) for $1 \leq p < \infty$ satisfies the properties (1.1) to (1.3) on the set $C^\uparrow$.

(1.1) $\sigma_p$ takes values only between 0 and 1, since $C^\uparrow$ is stochastically increasing and fulfils $0 \leq C^\uparrow - \Pi \leq C^\uparrow - \Pi$.

(1.2) $\sigma_p(C) = 0$ holds if and only if $C = \Pi$.

(1.3) Suppose $C = C^\uparrow$ is completely dependent. Then $\partial_1 C^\uparrow(u, v) \in \{0, 1\}$ by (3.15) and $\partial_1 C^\uparrow(u, v) = \mathbb{I}_{(0, 1]}(u)$ by Definition S.1.1(2). Thus, $C^\uparrow = C^\uparrow$ which yields $\sigma_p(C) = 1$. On the other hand, if $C$ is not completely dependent, then an analogous argument shows that $C^\uparrow < C^\uparrow$ on a set of positive measure such that

$$\sigma_p(C) = \frac{\|C^\uparrow - \Pi\|_p}{\|C^\uparrow - \Pi\|_p} < \frac{\|C^\uparrow - \Pi\|_p}{\|C^\uparrow - \Pi\|_p} = 1.$$

**Proof of the statements in Example 2.7** We introduce the concordance functional

$$Q(C_1, C_2) := 4 \int_{[0, 1]^2} C_1(u, v) \, dC_2(u, v) - 1.$$
and point out for later reference that $Q$ is symmetric and fulfils
\[ Q(C_1, C_2) \leq Q(C_1', C_2) \tag{3.16} \]
whenever $C_1 \leq C_1'$. Then the four measures of concordance (see Definition S.I.5) Spearman’s $\rho$, Kendall’s $\tau$, Gini’s $\gamma$ and Blomqvist’s $\beta$ are given by (see, e.g., Chapter 5 in Nelsen (2006))
\[
\begin{align*}
\rho(C) &= 3Q(C, \Pi) = 12 \int_{[0,1]^2} C(u, v) \, d\lambda(u, v) - 3 \\
\tau(C) &= Q(C, C) = 4 \int_{[0,1]^2} C(u, v) \, dC(u, v) - 1 \\
\gamma(C) &= Q(C, C^-) + Q(C, C^+) = 2 \int_{[0,1]^2} |u + v - 1| - |u - v| \, dC(u, v) \\
\beta(C) &= 4C \left( \frac{1}{2}, \frac{1}{2} \right) - 1,
\end{align*}
\]
First of all, $\beta$ does not satisfy (1.3) on $C^\uparrow$ because the copula$^5$
\[
C(u, v) = \begin{cases} 
2\Pi(u, v) & \text{if } (u, v) \in [0, 1/2]^2 \\
C^+(u, v) & \text{else}
\end{cases}
\]
is stochastically increasing with $C \neq C^+$, yet $\beta(C) = 4C(1/2, 1/2) - 1 = 1 = \beta(C^+)$. We now show that $\rho$, $\tau$ and $\gamma$ all satisfy the properties (1.1)–(1.3) on $C^\uparrow$. Since any concave function $f : [0, 1] \to [0, v]$ with $f(0) = 0$ and $f(1) = v$ satisfies $f(u) \geq uv = \Pi(u, v)$, any stochastically increasing copula $C$ satisfies
\[ \Pi \leq C = C^\uparrow \leq C^+ . \tag{3.17} \]
Hence we conclude from Definition S.I.5(4) that $0 = \kappa(\Pi) \leq \kappa(C^\uparrow) \leq \kappa(C^+) = 1$. It remains to verify properties (1.2) and (1.3) for $\rho$, $\tau$ and $\gamma$.
First, we look at Spearman’s $\rho$. By Proposition 2.8, $R_\rho$ coincides with $R_{\sigma}$, so that, in view of Example 2.6 with $p = 1$, the properties (1.2) and (1.3) hold.
Next, consider Kendall’s $\tau$. In order to prove (1.2), we assume $\tau(C) = \tau(\Pi)$, i.e. $Q(C, C) = Q(\Pi, \Pi)$, for some $C \in C^\uparrow$. In view of (3.16) and (3.17) we obtain $Q(\Pi, \Pi) \leq Q(C, \Pi) \leq Q(C, C) = Q(\Pi, \Pi)$ so that
\[
0 \leq 4 \int_{[0,1]^2} |C(u, v) - \Pi(u, v)| \, d\lambda(u, v) = 4 \int_{[0,1]^2} C(u, v) - \Pi(u, v) \, dC(u, v) = Q(C, \Pi) - Q(\Pi, \Pi) = 0
\]
which indeed implies $C = \Pi$. For the proof of (1.3), we suppose $\tau(C) = \tau(C^+)$, i.e. $Q(C, C) = Q(C^+, C^+)$. In view of (3.16) and (3.17) we obtain $Q(C, C) \leq Q(C, C^+) \leq Q(C^+, C^+) = Q(C, C)$ so that
\[
0 \leq 4 \int_0^1 |u - C(u, u)| \, du = 4 \int_0^1 u - C(u, u) \, du
\]
$^5$ $C$ is a so-called ordinal sum; see (Nelsen, 2006, Sect. 3.2.2).
Therefore \( C(u, u) = u \) for all \( u \in [0, 1] \) so that \( C = C^+ \) (see (Durante and Sempi, 2016, Ex 2.6.4)).  

Finally, we turn to Gini’s \( \gamma \). In order to prove (1.2), we assume \( \gamma(C) = \gamma(\Pi) \), i.e.

\[
Q(C, C^+) + Q(C, C^-) = Q(\Pi, C^+) + Q(\Pi, C^-),
\]

for some \( C \in C^+ \). In view of (3.16) and (3.17) we obtain

\[
Q(\Pi, C^+) + Q(\Pi, C^-) \leq Q(C, C^+) + Q(\Pi, C^-) \leq Q(C, C^-) + Q(\Pi, C^-) = Q(\Pi, C^+) + Q(\Pi, C^-)
\]

so that

\[
0 \leq 4 \int_0^1 |C(u, u) - \Pi(u, u)| \, du = 4 \int_0^1 C(u, u) - \Pi(u, u) \, du = Q(C, C^+) - Q(\Pi, C^+) = 0.
\]

It follows that \( C(u, u) = \Pi(u, u) \), and Proposition 2.1 in Durante and Papini (2009) yields \( C = \Pi \). For the proof of (1.3), we suppose \( \gamma(C) = \gamma(C^+) \), i.e.

\[
Q(C, C^+) + Q(C, C^-) = Q(C^+, C^+) + Q(C^+, C^-) = Q(C, C^+) + Q(C, C^-),
\]

In view of (3.16) and (3.17) we obtain

\[
Q(C, C^+) + Q(C, C^-) \leq Q(C^+, C^+) + Q(C^+, C^-) \leq Q(C^+, C^-) + Q(C, C^-) = Q(C, C^+) + Q(C, C^-),
\]

which implies

\[
0 \leq 4 \int_0^1 |u - C(u, u)| \, du = 4 \int_0^1 u - C(u, u) \, du
\]

\[
= 4 \int_{[0,1]^2} C^+(u, v) - C(u, v) \, dC^+(u, v) = Q(C^+, C^+) - Q(C, C^+) = 0.
\]

Therefore \( C(u, u) = u \) for all \( u \in [0, 1] \) so that \( C = C^+ \) (Durante and Sempi, 2016, Ex. 2.6.4).

**Proof of Proposition 2.8**  This follows readily from the fact that \( C^\uparrow \geq \Pi \) since

\[
R_{\sigma_1}(C) = \frac{\|C^\uparrow - \Pi\|_1}{\|C^+ - \Pi\|_1} = 12 \int_{[0,1]^2} C^\uparrow(u, v) - uv \, dI(u, v)
\]

\[
= 12 \int_{[0,1]^2} C^\uparrow(u, v) \, dI(u, v) - 3 = R_\rho(C).
\]

**Proof of Theorem 2.11**  In view of Definition S.1.5, we have \( \kappa(C^\downarrow) = \kappa(C^- * C^\uparrow) = -\kappa(C^\uparrow) \). Consequently, we know from Lemma A.1 and the monotonicity of \( \kappa \) with respect to the pointwise ordering that

\[
-\kappa(C^\uparrow) = \kappa(C^\downarrow) \leq \kappa(C) \leq \kappa(C^\uparrow),
\]

6The observation that \( \tau(C) = \tau(C^+) \) implies \( C = C^+ \) also in the multivariate case is contained in (Fuchs, McCord and Schmidt, 2018, Thm. 3.2).
Rearranged dependence measures

which implies $|\kappa(C)| \leq \kappa(C^\dagger) = R_\kappa(C)$. Moreover, if $C$ is stochastically monotone we have $C = C^\dagger$ or $C = C^{\dagger}$ and, therefore, $|\kappa(C)| = \kappa(C^\dagger)$.

**Proof of Proposition 2.13** First, we point out that that the Markov product of two copulas $C$ and $D$ satisfies

$$\partial_1 (C \ast D)(\cdot, v) = \partial_1 \int_0^1 \partial_2 C(\cdot, t) \cdot \partial_1 D(t, v) \; dt \leq \partial_1 D(\cdot, v) \tag{3.18}$$

for all $v \in [0, 1]$, where “≤” denotes the majorization order introduced in Definition S.1.7. This follows from Theorem S.1.8(3) and the fact that $\partial_1 (C \ast D)(u, v) = T_C \partial_1 D(\cdot, v)(u)$. In particular,

$$(C \ast D)^\dagger(u, v) \leq D^\dagger(u, v).$$

Now suppose $X, Y$ and $Z$ are continuous random variables such that $Y$ and $Z$ are conditionally independent given $X$. Then $C_{ZY} = C_{ZX} \ast C_{XY}$ in view of Theorem 3.1 in Darsow, Nguyen and Olsen (1992), and (3.18) yields

$$C_{ZY}^\dagger = (C_{ZX} \ast C_{XY})^\dagger \leq C_{XY}^\dagger.$$

Thus, the data processing inequality $R_\mu(C_{ZY}) = \mu(C_{ZY}^\dagger) \leq \mu(C_{XY}^\dagger) = R_\mu(C_{XY})$ follows from the monotonicity of $\mu$.

**Proof of Corollary 2.14** The data processing inequality in Proposition 2.13 states that

$$R_\mu(f(X), Y) \leq R_\mu(X, Y)$$

for all measurable functions $f$. If, in addition, $X$ and $Y$ are independent given $f(X)$, a second application of Proposition 2.13 yields $R_\mu(X, Y) \leq R_\mu(f(X), Y)$, and equality holds.

**Appendix B: Proofs of the results in Section 3**

**Proof of Theorem 3.1** The equality $C_{N_1, N_2}^\# (A^\dagger) = C_{N_1, N_2}^\# (A^{\dagger})$ follows directly from the definition of Algorithm 1 and the characterization (3.6). It remains to show that the matrix $A^\dagger$ satisfies indeed the properties in (3.2). To do so, we calculate

$$\sum_{\ell=1}^{N_2} a_{k\ell} = \sum_{\ell=1}^{N_2} \hat{B}_{k\ell} - \hat{B}_{k\ell}^{-1} = \hat{B}_{k}^{N_2} - \hat{B}_{k}^{0} = \hat{B}_{k}^{N_2} - \sum_{\ell=1}^{N_2} a_{k\ell} = N_2$$

as well as

$$\sum_{k=1}^{N_1} a_{kj} = \sum_{k=1}^{N_1} \hat{B}_{k}^{-1} - \hat{B}_{k}^{-1} = \sum_{k=1}^{N_1} \hat{B}_{k}^{-1} - \hat{B}_{k}^{-1} = \sum_{j=1}^{N_1} \sum_{k=1}^{N_1} a_{k\ell} - \sum_{j=1}^{N_1} \sum_{k=1}^{N_1} a_{kj} = \ell N_1 - (\ell - 1)N_1 = N_1.$$

The nonnegativity of $a_{k\ell}^\dagger$ follows by construction.

**Proof of Theorem 3.2** We will start by showing a contraction property of the (SI)-rearrangement with respect to $D_\mu$. For all copulas $C$ and $D$, it holds by Theorem S.1.8(b)

$$\partial_1 C^\dagger(\cdot, v) - \partial_1 D^\dagger(\cdot, v) \leq \partial_1 C(\cdot, v) - \partial_1 D(\cdot, v)$$
for all \( v \) in \([0, 1]\), where \( \preceq \) denotes the majorization order introduced in Definition S.I.7. Thus, due to Theorem S.I.8, we have for all \( v \in [0, 1) \) and any \( 1 \leq p < \infty \)

\[
\int_0^1 |\partial_1 C(u, v) - \partial_1 D(u, v)|^p \, du \leq \int_0^1 |\partial_1 C(u, v) - \partial_1 D(u, v)|^p \, du .
\]

and integrating with respect to \( v \) yields \( D_p(C^\top, D^\top) \leq D_p(C, D) \). Now it follows by similar arguments as in the proof of Theorem 4.5.8 in Durante and Sempi (2016) (these authors considered the case \( N_1 = N_2 \)) that

\[
0 \leq D_p(C_{N_1,N_2}^\#(C)^\top, C^\top) \leq D_p(C_{N_1,N_2}^\#(C), C) \to 0 .
\]

**Proof of Theorem 3.4** The almost sure convergence of \( D_1(\hat{C}_{N_1,N_2,n}^\#, C) \to 0 \) follows from Theorem 3.12 in Junker, Griessenberger and Trutschnig (2021), where \( \hat{C}_{N_1,N_2,n}^\# \) is a genuine copula. Thus, an application of the continuity property given in Theorem 3.2 implies

\[
0 \leq D_1((\hat{C}_{N_1,N_2,n}^\#)^\top, C^\top) \leq D_1(\hat{C}_{N_1,N_2,n}^\#, C) \to 0 .
\]

and therefore \( \hat{R}_\mu \to R_\mu(C) \) almost surely.

**Supplementary Material**

**Supplement to “Rearranged dependence measures”**

This supplement contains basic facts about copulas and monotone rearrangements and provides proofs of the multivariate results of Section 2.4.

**Funding**

C. Strothmann gratefully acknowledges financial support from the German Academic Scholarship Foundation. The work of H. Dette was supported by the DFG Research Unit 5381 Mathematical Statistics in the Information Age. The authors are grateful to two referees for their constructive comments on an earlier version of this paper.

**References**

ANEVSKI, D. and FOUGÈRES, A.-L. (2019). Limit properties of the rearrangement for density and regression function estimation. *Bernoulli* 25 549 – 583. MR3892329

ANSARI, J. and RÜSCHENDORF, L. (2021). Sklar’s theorem, copula products, and ordering results in factor models. *Depend. Model.* 9 267–306. MR4327840

AUDHY, A., DEB, N. and NANDY, S. (2021). Exact Detection Thresholds for Chatterjee’s Correlation. https://arxiv.org/abs/2104.15140.

BENJAMINI, Y. and HOCHBERG, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. Roy. Statist. Soc. Ser. B* 57 289-300. MR1325392

BENNETT, C. and SHARPLEY, R. C. (1988). *Interpolation of Operators.* Academic Press, Boston. MR0928802

BERGSMA, W. and DASSIOS, A. (2014). A consistent test of independence based on a sign covariance related to Kendall’s tau. *Bernoulli* 20 1006 – 1028. MR3178526

BLUM, J. R., KIEFER, J. and ROSENBLATT, M. (1961). Distribution Free Tests of Independence Based on the Sample Distribution Function. *Ann. Math. Statist.* 32 485 – 498. MR0125690
Rearranged dependence measures

CAMIRAND-LEMYRE, F., CARROLL, R. J. and DELAIGLE, A. (2022). Semiparametric estimation of the distribution of episodically consumed foods measured with error. J. Amer. Statist. Assoc. 117 469–481. MR4399099

CAO, S. and BICKEL, P. J. (2020). Correlations with tailored extremal properties. http://arxiv.org/abs/2008.10177.

CHATTERJEE, S. (2021). A New Coefficient of Correlation. J. Amer. Statist. Assoc. 116 2009–2022. MR4353729

CHEN, S. X. and HUANG, T. M. (2007). Nonparametric Estimation of Copula Functions for Dependence Modelling. Canad. J. Statist. 35 265–282. MR2393609

CHERNOZHUKOV, V., FERNANDEZ-VAL, I. and GALICHON, A. (2009). Improving point and interval estimators of monotone functions by rearrangement. Biometrika 96 559–575. MR2538757

CHERNOZHUKOV, V., FERNANDEZ-VAL, I. and GALICHON, A. (2010). Quantile and Probability Curves Without Crossing. Econometrica 78 1093–1125. MR2667913

CHONG, K. M. and RICE, N. M. (1971). Equimeasurable Rearrangements of Functions. Queen’s University, Kingston. MR0372140

COVER, T. M. and THOMAS, J. A. (2006). Elements of Information Theory, 2nd ed. Wiley-Interscience, Hoboken. MR2239987

CSÖRGO, S. (1985). Testing for independence by the empirical characteristic function. J. Multivariate Anal. 16 290–299. MR0793494

DARSOW, W. F., NGUYEN, B. and OLSEN, E. T. (1992). Copulas and Markov processes. Illinois J. Math. 36 600–642. MR1215798

DEB, N., GHOSAL, P. and SEN, B. (2020). Measuring association on topological spaces using kernels and geometric graphs. http://arxiv.org/abs/2010.01768.

DETTE, H., NEUMEYER, N. and PILZ, K. F. (2006). A simple nonparametric estimator of a strictly monotone regression function. Bernoulli 12 469–490. MR2227277

DETTE, H., SIBURG, K. F. and STOIMENOV, P. A. (2013). A copula-based non-parametric measure of regression dependence. Scand. J. Stat. 40 21–41. MR3024030

DETTE, H. and VOLGUSHEV, S. (2008). Non-crossing non-parametric estimates of quantile curves. J. R. Stat. Soc. Ser. B Stat. Methodol. 70 609–627. MR2420417

DETTE, H. and WU, W. (2019). Detecting relevant changes in the mean of nonstationary processes, a mass excess approach. Ann. Statist. 47 3578–3608. MR4025752

DURANTE, F. and PAPINI, P. L. (2009). Componentwise concave copulas and their asymmetry. Kybernetika (Prague) 45 1003–1011. MR2650079

DURANTE, F. and SEMPI, C. (2016). Principles of Copula Theory. CRC Press, Boca Raton. MR3443023

FERMANIAN, J. D., RADULOVIC, D. and WEGKAMP, M. (2004). Weak convergence of empirical copula processes. Bernoulli 10 847–860. MR2093613

FUHS, S., MCCORD, Y. and SCHMIDT, K. D. (2018). Characterizations of copulas attaining the bounds of multivariate Kendall’s Tau. J. Optim. Theory Appl. 175 424–438. MR3825632

GAMBOA, F., GREMAUD, P., KLEIN, T. and LAGNOUX, A. (2022). Global sensitivity analysis: A novel generation of might estimators based on rank statistics. Bernoulli 28 2345–2374. MR4474546

GEENENS, G. and LAFAYE DE MICHEAUX, P. (2022). The Hellinger Correlation. J. Amer. Statist. Assoc. 117 639–653. MR4436302

GENEST, C., NEŠLEHOVÁ, J. G. and REMILLARD, B. (2017). Asymptotic behavior of the empirical multilinear copula process under broad conditions. J. Multivariate Anal. 159 82–110. MR3668549

GRETTON, A., FUKUMIZU, K., TEO, C., SONG, L., SCHOLKOPF, B. and SMOLA, A. (2008). A Kernel Statistical Test of Independence. In Advances in Neural Information Processing Systems (J. PLATT, D. KOLLER, Y. SINGER and S. ROWEIS, eds.) 20. Curran Associates, Inc.

GRIESSENBERGER, F., JUNKER, R. R. and TRUTSCHING, W. (2022). On a multivariate copula-based dependence measure and its estimation. Electron. J. Stat. 16. MR4401220

HARDY, G. H., LITTLEWOOD, J. E. and POLYA, G. (1988). Inequalities. Cambridge Mathematical Library. Cambridge University Press, Cambridge. Reprint of the 1952 edition. MR944909

HOFERT, M., KOJADINOVIC, I., MÄCHLER, M. and YAN, J. (2020). copula: Multivariate dependence with copulas R package version 1.0-1 available at https://CRAN.R-project.org/package=copula.

JUNKER, R. R., GRIESSENBERGER, F. and TRUTSCHING, W. (2021). Estimating scale-invariant directed dependence of bivariate distributions. Comput. Statist. Data Anal. 153 107058. MR4141460
KASPER, T., GRIESSENBERGER, F., JUNKER, R. R., PETZEL, V. and TRUTSCHNIG, W. (2022). qad: Quantification of Asymmetric Dependence R package version 1.0.4.

KINNEY, J. B. and ATWAL, G. S. (2014). Equitability, mutual information, and the maximal information coefficient. Proc. Natl. Acad. Sci. USA 111 3354–3359. MR3200177

LEHMANN, E. L. (1959). Testing Statistical Hypotheses. Wiley, New York. MR0107933

LI, X., MIKUSINSKI, P. and TAYLOR, M. D. (1998). Strong Approximation of Copulas. J. Math. Anal. Appl. 225 608–623. MR1644300

LI, X., MIKUSINSKI, P., SHERWOOD, H. and TAYLOR, M. D. (1997). On approximation of copulas. In Distributions with given Marginals and Moment Problems (V. Beneš and J. Štěpán, eds.) 107–116. Springer, Dordrecht. MR1614663

LIN, Z. and HAN, F. (2022). On boosting the power of Chatterjee’s rank correlation. To appear in: Biometrika.

MARSHALL, A. W., OLKIN, I. and ARNOLD, B. C. (2011). Inequalities: Theory of Majorization and Its Applications, 2nd ed. Springer Series in Statistics. Springer, New York. MR2759813

MIKUSINSKI, P., SHERWOOD, H. and TAYLOR, M. (1992). Shuffles of min. Stochastica 13 61–74. MR1197328

NELSEN, R. B. (2006). An Introduction to Copulas, 2nd ed. Springer Series in Statistics. Springer, New York. MR2197664

OMELKA, M., GJIBELS, I. and VERAVERBEKE, N. (2009). Improved kernel estimation of copulas: Weak convergence and goodness-of-fit testing. Ann. Statist. 37 3023 – 3058. MR2541454

RESHEF, D. N., RESHEF, Y. A., FINUCANE, H. K., GROSSMAN, S. R., MCEVAN, G., TURNBAUGH, P. J., LANDER, E. S., MITZENMACHER, M. and SABETI, P. C. (2011). Detecting novel associations in large data sets. Science 334 1518–1524.

ROSENBLATT, M. (1975). A Quadratic Measure of Deviation of Two-Dimensional Density Estimates and A Test of Independence. Ann. Statist. 3 1 – 14. MR0428579

RYFF, J. V. (1965). Orbits of $L^1$-functions under doubly stochastic transformations. Trans. Amer. Math. Soc. 117 92–100. MR0209866

RYFF, J. V. (1970). Measure preserving transformations and rearrangements. J. Math. Anal. Appl. 31 449–458. MR0419734

SCHWEIZER, B. and WOLFF, E. F. (1981). On nonparametric measures of dependence for random variables. Ann. Statist. 9 879–885. MR0619291

SHI, H., DRTON, M. and HAN, F. (2021). On Azadkia-Chatterjee’s conditional dependence coefficient. http://arxiv.org/abs/2108.06827.

SHI, H., DRTON, M. and HAN, F. (2022). On the power of Chatterjee’s rank correlation. Biometrika 109 317–333. MR4430960

SIBURG, K. F. and STROTHMANN, C. (2021). Stochastic monotonicity and the Markov product for copulas. J. Math. Anal. Appl. 503 125348. MR4263102

SPELLMAN, P. T., GAVIN, S., ZHANG, M. Q., IYER, V. R., ANDERS, K., EISEN, M. B., BROWN, P. O., BOSTSTEIN, D. and FUTCHER, B. (1998). Comprehensive Identification of Cell Cycle-regulated Genes of the Yeast Saccharomyces cerevisiae by Microarray Hybridization. Mol. Biol. Cell 9 3273–3297.

STONE, C. J. (1984). An Asymptotically Optimal Window Selection Rule for Kernel Density Estimates. The Annals of Statistics 12 1285–1297. MR0760688

STROTHMANN, C., DETTE, H. and SIBURG, K. F. (2022). Online supplement to: Rearranged dependence measures. http://arxiv.org/abs/2201.03329.

SZÉKELY, G. J., RIZZO, M. L. and BÁKIROV, N. K. (2007). Measuring and testing dependence by correlation of distances. Ann. Statist. 35 2769 – 2794. MR2382665

R CORE TEAM (2021). R: A language and environment for statistical computing R Foundation for Statistical Computing, Vienna.

TRUTSCHNIG, W. (2011). On a strong metric on the space of copulas and its induced dependence measure. J. Math. Anal. Appl. 384 690–705. MR2825218

ZHANG, K. (2019). BET on Independence. J. Amer. Statist. Assoc. 114 1620–1637. MR4047288
Online supplement to: Rearranged dependence measures

CHRISTOPHER STROTHMANN\textsuperscript{1,a}, HOLGER DETTE\textsuperscript{2,c} and KARL FRIEDRICH SIBURG\textsuperscript{1,b}

\textsuperscript{1}Department of Mathematics, TU Dortmund University, Vogelpothsweg 87, 44221 Dortmund, Germany, \textsuperscript{a}christopher.strothmann@mathematik.tu-dortmund.de, \textsuperscript{b}karl.f.siburg@mathematik.tu-dortmund.de

\textsuperscript{2}Department of Mathematics, Ruhr-University Bochum, Universitätsstraße 150, 44780 Bochum, Germany, \textsuperscript{c}holger.dette@rub.de

This online supplement to Strothmann, Dette and Siburg (2022) contains two parts. S.I collects some basic facts about copulas as well as monotone rearrangements. S.II presents the proofs of the multivariate results of Section 2.4.

Supplement S.I: Results from the literature

In this section, we present some basic facts about copulas and monotone rearrangements, which will be frequently used throughout the proofs of our results in Appendix A and B. We start with the definition of a bivariate copula, which is a distribution function on the unit square with uniform univariate margins.

**Definition S.I.1.** A function $C : [0,1]^2 \rightarrow [0,1]$ is called a (bivariate) copula if

1. $C$ is grounded, i.e. $C(0,v) = C(u,0) = 0$ for all $u,v \in [0,1]$
2. $C$ has uniform margins, i.e. $C(1,u) = C(u,1) = u$ for all $u \in [0,1]$
3. $C$ is 2-increasing, i.e. the $C$-volume of every rectangle $R = [u_1,u_2] \times [v_1,v_2]$ is nonnegative:

$$V_C(R) := C(u_2,v_2) - C(u_1,v_2) - C(u_2,v_1) + C(u_1,v_1) \geq 0.$$ 

The set of all copulas is denoted by $C$. We refer to the lower Fréchet-Hoeffding bound by $C^- (u,v) := \max\{u + v - 1, 0\}$, to the independence (or product) copula by $\Pi(u,v) := uv$, and to the upper Fréchet-Hoeffding bound by $C^+ (u,v) := \min\{u,v\}$. Any copula $C$ satisfies $C^- \leq C \leq C^+$.

**Definition S.I.2.** The Markov product of two copulas $C$ and $D$ is defined as the copula

$$(C * D)(u,v) := \int_0^1 \partial_2 C(u,t) \partial_1 D(t,v) \, dt.$$ 

A comprehensive review of the Markov product can be found in Durante and Sempi (2016).

**Definition S.I.3.** A linear operator $T : L^1([0,1]) \rightarrow L^1([0,1])$ is called a Markov operator if

1. $T$ is positive, i.e. $Tf \geq 0$ whenever $f \geq 0$
2. $T1_{[0,1]} = 1_{[0,1]}$
3. $T$ preserves the integral, i.e. $\int_0^1 Tf(t) \, dt = \int_0^1 f(t) \, dt$ for all $f \in L^1([0,1])$. 

1
The following result shows that copulas and Markov operators are closely linked and that the composition of Markov operators corresponds to the Markov product of copulas. A proof can be found in Olsen, Darsow and Nguyen (1996).

**Theorem S.I.4.** Let $C$ be a copula and $T$ be a Markov operator. Then

$$C_T(u,v) := \int_0^u T\mathbb{1}_{[0,v]}(t) \, dt \text{ and } T_C f(u) := \partial_u \int_0^1 \partial_2 C(u,v) f(v) \, dv$$

define a copula $C_T$ and a Markov operator $T_C$, respectively. The correspondence $C \mapsto T_C$ is bijective with $T_{C_T} = T$ and $C_{T_C} = C$. Moreover,

$$T_{C_1 \ast C_2} = T_{C_1} \circ T_{C_2}$$

holds for all copulas $C_1$ and $C_2$.

The following definition of a concordance measure is adapted from Durante and Sempi (2016).

**Definition S.I.5.** A function $\kappa : C \to [-1, 1]$ is called a measure of concordance if

1. $\kappa(C^-) = -1$, $\kappa(\Pi) = 0$ and $\kappa(C^+) = 1$
2. $\kappa(C^\top) = \kappa(C)$, where $C^\top(u,v) := C(v,u)$
3. $\kappa(C^\ast C) = \kappa(C \ast C^-) = -\kappa(C)$
4. $\kappa$ is monotone w.r.t. the pointwise order on the set of copulas
5. $\kappa$ is continuous w.r.t. the pointwise\(^1\) convergence of copulas.

For the decreasing rearrangement $f^* : [0, 1] \to \mathbb{R}$ of a measurable function $f : [0, 1]^d \to \mathbb{R}$, we state the following properties.

**Proposition S.I.6.** For any two measurable functions $f$ and $g$ on the measure space $([0, 1]^d, m)$, the following assertions hold:

1. $f^*$ is decreasing and right-continuous on $[0, 1]$.
2. $f \leq g$ implies $f^* \leq g^*$.
3. There exists an $m$-$\lambda$-preserving transformation $\sigma : [0, 1]^d \to [0, 1]$ such that $f = f^* \circ \sigma$.
4. The decreasing rearrangement is $L^p$-invariant for $1 \leq p \leq \infty$, i.e.

$$\|f\|_p = \|f^*\|_p .$$

**Proof.** Property (1) is stated in Theorem 4.2, properties (2) and (4) can be found in Proposition 4.3, and property (3) is stated in Theorem 6.2 of Chong and Rice (1971).

Closely linked to the decreasing rearrangement of measurable functions is an ordering widely known as the majorization order, introduced by Hardy, Littlewood and Pólya for vectors, and by Ryff (1965) for functions.

\(^1\)As copulas are continuous function on a compact set, pointwise and uniform convergence are equivalent.
**Definition S.I.7.** Suppose $f, g \in L^1([0,1], \lambda)$. Then $f$ is majorized by $g$, denoted by $f \preceq g$, if
\[
\int_0^t f^*(s) \, ds \leq \int_0^t g^*(s) \, ds
\]
holds for all $t \in [0,1]$, as well as
\[
\int_0^1 f^*(s) \, ds = \int_0^1 g^*(s) \, ds.
\]

**Theorem S.I.8.** For $f, g \in L^1([0,1], \lambda)$, the following statements are equivalent:

1. $f$ is majorized by $g$, i.e. $f \preceq g$.
2. For every convex function $\phi : \mathbb{R} \to \mathbb{R}$ we have
\[
\int_0^1 \phi(f(s)) \, ds \leq \int_0^1 \phi(g(s)) \, ds.
\]
3. There exists a Markov operator $T$ such that $f = Tg$.

Furthermore, the following inequalities hold:

(a)
\[
\int_0^1 |f^*(s)g^*(1-s)| \, ds \leq \int_0^1 |f(s)g(s)| \, ds \leq \int_0^1 |f^*(s)g^*(s)| \, ds.
\]

(b)
\[
f^* - g^* \preceq f - g.
\]

**Proof.** The equivalence of (1) and (3) is shown in (Day, 1973, Thm. 4.9), while that of (1) and (2) is contained in (Chong, 1974, Thm. 2.5). The proofs of (a), called the Hardy-Littlewood inequality, and (b) can be found in (Day, 1972, (6.2) and (6.1)).

**Supplement S.II: Remaining proofs of the results in Section 2.4**

**Proof of Equation (2.10).** In order to show that the stochastically increasing rearrangement $C^\uparrow$ is indeed a copula, we verify the properties (1) to (3) of Definition S.I.1.

1. It follows from $(K_C)^*(u,0) = 0^* = 0$ that $C^\uparrow(u,0) = 0$. The identity $C^\uparrow(0,v) = 0$ is trivial by Equation (2.9).
2. By definition, we have
\[
C^\uparrow(u,1) = \int_0^u (K_C)^*(s,1) \, ds = \int_0^u 1^* \, ds = u.
\]
In view of Proposition S.I.6(3), we further obtain that
\[
C^\uparrow (1, v) = \int_0^1 (K_C)^\uparrow (s, v) \, ds = \int_{[0,1]^d} (K_C)^\uparrow (\sigma_v(s), v) \, d\mu_{C^\downarrow} (s) = \int_{[0,1]^d} K_C (s, [0,v]) \, d\mu_{C^\downarrow} (s) = C(1, \ldots, 1, v) = v.
\]

3. Since \( K_C (s, \cdot) \) is a probability measure for each \( s \), we see that \( 0 \leq K_C (\cdot, [0,v_1]) \leq K_C (\cdot, [0,v_2]) \) whenever \( v_1 \leq v_2 \). Combining this with Proposition S.I.6(2) yields \( (K_C)^\uparrow (\cdot, v_1) \leq (K_C)^\uparrow (\cdot, v_2) \).

Thus, the \( C^\uparrow \)-volume of a rectangle \([u_1, u_2] \times [v_1, v_2] \) satisfies
\[
V_{C^\uparrow} ([u_1, u_2] \times [v_1, v_2]) = C^\uparrow (u_2, v_2) - C^\uparrow (u_1, v_2) - C^\uparrow (u_2, v_1) + C^\uparrow (u_1, v_1) = \int_{u_1}^{u_2} (K_C)^\uparrow (s, v_2) - (K_C)^\uparrow (s, v_1) \, ds \geq 0.
\]

Thus \( C^\uparrow \) is a bivariate copula. Moreover, since \( s \mapsto (K_C)^\uparrow (s, v) \) is decreasing, \( C^\uparrow \) is stochastically increasing.

**Proof of (M 1.1)–(M 1.3).** We will now prove properties (M 1.1)–(M 1.3) for \( R_{\mu} (C) = \mu (C^\uparrow) \). For this, we say that the copula \( C \) is completely dependent if there exists a measurable function \( f \) such that \( V = f (U) \). It is proven in Griessenberger, Junker and Trutschnig (2022) that \( C \) is completely dependent if, and only if, there is a \( \mu_{C^\downarrow} \)-\( \lambda \)-preserving transformation \( h : [0,1]^d \rightarrow [0,1] \) such that
\[
K_C (s, F) = \mathbb{I}_F (h(s)) \quad (1)
\]
for all Borel measurable sets \( F \).

(M1.1) Since \( \mu \) only takes values between 0 and 1, we obtain the first assertion.

(M1.2) Suppose \( U \) and \( V \) are independent. Then, following the proof of Theorem 5.6 in Griessenberger, Junker and Trutschnig (2022), \( K_C (s, [0,v]) = v \). Thus, \( (K_C)^\uparrow (s, v) = v, C^\uparrow = \Pi \) and \( \mu (C^\uparrow) = \mu (\Pi) = 0 \). If, on the other hand, \( \mu (C^\uparrow) = 0 \), we conclude \( C^\uparrow = \Pi \) by the properties of \( \mu \). But then, \( (K_C)^\uparrow (s, v) = v \) and \( K_C (s, [0,v]) = (K_C)^\uparrow (\sigma_v (s), v) = v \) for some measure-preserving transformation \( \sigma_v \). Therefore, \( C (u, v) = v C^\uparrow (u) \) and \( U \) and \( V \) are independent.

(M1.3) If \( C \) is completely dependent, then \( C^\uparrow = C^\uparrow^* \) and \( \mu (C^\uparrow) = \mu (C^\uparrow^*) = 1 \) by definition. On the other hand, \( \mu (C^\uparrow) = 1 \) implies \( C^\uparrow = C^\uparrow^* \) by the properties of \( \mu \). Thus, \( K_C (s, [0,v]) = \partial_1 C^\uparrow (\sigma_v (s), v) = \mathbb{I}_{[0,v]} (\sigma_v (s)) \), so \( C \) is completely dependent by (1).

**Proof of the information gain inequality.** To shorten the necessary notation, we consider the case of three random variables \( X_1, X_2 \) and \( Y \). The copulas of \( (X_1, Y) \), \( (X_1, X_2, Y) \) and \( (X_1, X_2) \) are denoted by \( D, C \) and \( C^{12} \), respectively. Following the proof of Theorem 5.6(5) in Griessenberger, Junker and Trutschnig (2022), we have by the disintegration theorem
\[
K_D (u_1, [0,v]) = \int_0^1 K_C (u_1, u_2, [0,v]) \, K_{C^{12}} (u_1, du_2),
\]
where \( K_C \) denotes the disintegration with respect to the first two variables as above. Since \( K_{C^{12}} (u_1, \cdot) \) is a probability measure for all \( u_1 \in [0,1] \), we have via Jensen’s inequality for any convex function \( \phi \)
\[
\int_0^1 \phi(K_D (u_1, [0,v])) \, du_1 = \int_0^1 \phi \left( \int_0^1 K_C (u_1, u_2, [0,v]) \, K_{C^{12}} (u_1, du_2) \right) \, du_1
\]
\[
\leq \int_0^1 \int_0^1 \phi \left( K_C(u_1, u_2, [0, v]) \right) K_{C_{12}}(u_1, du_2) \, du_1 \\
= \int_{[0,1]^2} \phi \left( K_C(u_1, u_2, [0, v]) \right) \, d\mu_{C_{12}}(u_1, u_2).
\]

By Theorem 2.5 of Chong (1974), we have \( K_D(\cdot, [0, v]) \preceq K_C(\cdot, [0, v]) \) for all \( v \in [0, 1] \) and therefore \( D^\uparrow \leq C^\uparrow \).

**Proof of the conditional independence property.** Suppose \( X_2, \ldots, X_d \) and \( Y \) are conditionally independent given \( X_1 \). Using the proof of Proposition 5.9 in Griessenberger, Junker and Trutschnig (2022), we have

\[
K_C(u_1, \ldots, u_d, [0, v]) = K_D(u_1, [0, v]),
\]

where \( D \) denotes the copula of \( X_1 \) and \( Y \) as above. Thus, similarly to the proof of the information gain inequality, we have for any convex function \( \phi \)

\[
\int_{[0,1]} \phi \left( K_D(u_1, [0, v]) \right) \, du_1 = \int_{[0,1]^d} \phi \left( K_D(u_1, [0, v]) \right) \, d\mu_{C_{1\ldots d}}(u_1, \ldots, u_d)
\]

\[
= \int_{[0,1]^d} \phi \left( K_C(u_1, \ldots, u_d, [0, v]) \right) \, d\mu_{C_{1\ldots d}}(u_1, \ldots, u_d).
\]

and therefore \( K_D^*(\cdot, [0, v]) = K_C^*(\cdot, [0, v]) \) and \( D^\uparrow = C^\uparrow \).

**References**

CHONG, K. M. (1974). Some extensions of a theorem of Hardy, Littlewood and Pólya and their applications. *Canadian J. Math.* 26 1321–1340. MR0352377

CHONG, K. M. and RICE, N. M. (1971). *Equimeasurable Rearrangements of Functions.* Queen’s University, Kingston. MR0372140

DAY, P. W. (1972). Rearrangement inequalities. *Canadian J. Math.* 24 930–943. MR0310156

DAY, P. W. (1973). Decreasing rearrangements and doubly stochastic operators. *Trans. Amer. Math. Soc.* 178 383–383. MR0318962

DURANTE, F. and SEMPI, C. (2016). *Principles of Copula Theory.* CRC Press, Boca Raton. MR3443023

GRIESENBERGER, F., JUNKER, R. R. and TRUTSCHNIG, W. (2022). On a multivariate copula-based dependence measure and its estimation. *Electron. J. Stat.* 16. MR4401220

OLSEN, E. T., DARSOW, W. F. and NGUYEN, B. (1996). Copulas and Markov operators. In *Distributions with fixed marginals and related topics* 244–259. Institute of Mathematical Statistics, Hayward. MR1485536

RYFF, J. V. (1965). Orbits of \( L_1 \)-functions under doubly stochastic transformations. *Trans. Amer. Math. Soc.* 117 92–100. MR0209866

STROTHMANN, C., DETTE, H. and SIBURG, K. F. (2022). Rearranged dependence measures. http://arxiv.org/abs/2201.03329.