

# A MEASURE OF MUTUAL COMPLETE DEPENDENCE

BY KARL FRIEDRICH SIBURG AND PAVEL A. STOIMENOV

*University of Dortmund, Germany*

Two random variables  $X$  and  $Y$  are mutually completely dependent (m.c.d.) if there is a measurable bijection  $f$  with  $P(Y = f(X)) = 1$ . For continuous  $X$  and  $Y$ , a natural approach to constructing a measure of dependence is via the distance between the copula of  $X$  and  $Y$  and the independence copula. We show that this approach depends crucially on the choice of the distance function. For example, the  $L^p$ -distances, suggested by Schweizer and Wolff, cannot generate a measure of (mutual complete) dependence, since every copula is the uniform limit of copulas linking m.c.d. variables.

Instead, we propose to use a modified Sobolev norm, with respect to which, mutual complete dependence cannot approximate any other kind of dependence. This Sobolev norm yields the first nonparametric measure of dependence capturing precisely the two extremes of dependence, i.e., it equals 0 if and only if  $X$  and  $Y$  are independent, and 1 if and only if  $X$  and  $Y$  are m.c.d.

**1. Introduction.** Let  $X$  and  $Y$  be two random variables.  $Y$  is defined (see [5]) to be completely dependent on  $X$  if there exists a Borel measurable function  $f$  such that

$$(1.1) \quad P(Y = f(X)) = 1.$$

$X$  and  $Y$  are called *mutually completely dependent* (m.c.d.) if  $Y$  is completely dependent on  $X$ , and  $X$  is completely dependent on  $Y$ . In other words,  $X$  and  $Y$  are m.c.d. if, and only if, there is a Borel measurable bijection  $f$  satisfying (1.1).

Stochastic independence and mutual complete dependence are exactly opposite in character. The former case entails complete unpredictability of either random variable from the other, whereas the latter

---

*AMS 2000 subject classifications:* Primary 62E10; secondary 62H20.

*Keywords and phrases:* Nonparametric measures of dependence, mutual complete dependence, copulas, Sobolev norm.

corresponds to complete predictability. Consequently, it is a natural desideratum for a measure of dependence between  $X$  and  $Y$  to measure the strength of mutual complete dependence, with extreme values of 0 if and only if  $X$  and  $Y$  are independent, and 1 if and only if  $X$  and  $Y$  are m.c.d.

However, even the maximal correlation, introduced by Gebelein [3] and considered by Rényi [8] in many respects superior to other measures of dependence, fails to capture the extremes of stochastic dependence. The maximal correlation is defined by

$$(1.2) \quad \tilde{\rho}(X, Y) = \sup_{g, h} \rho(g(X), h(Y)),$$

where the supremum is taken over all Borel measurable functions  $g, h$  such that  $\text{Var } g(X), \text{Var } h(Y) \in (0, \infty)$ , and  $\rho$  denotes Pearson's correlation coefficient. In fact, besides computational difficulties,  $\tilde{\rho}$  equals 1 too often since two random variables which are m.c.d. have maximum correlation 1, but the converse is not true.

The lack of quantities which measure mutual complete dependence could be attributed to the fact that mutual complete dependence seems incompatible with the concept of weak convergence of bivariate distribution functions. In particular, it has been shown in [4] that it is possible to construct a sequence of pairs of m.c.d. random variables, all having uniform margins, which converges in law to a pair of independent random variables. This disturbing and counterintuitive phenomenon has led to the concept of monotone dependence, which corresponds to mutual complete dependence when in (1.1) the class of Borel measurable functions  $f$  is restricted to a.s. monotone ones. Kimeldorf and Sampson [4] argue that monotone dependence could be interpreted as the opposite of stochastic independence because the property of monotone dependence is preserved under weak convergence. They introduce the monotone correlation, given by

$$(1.3) \quad \rho^*(X, Y) = \sup_{g, h} \rho(g(X), h(Y)),$$

where the supremum is taken only over monotone functions. As noted in [4], however, two random variables which are monotone dependent have monotone correlation 1 but, again, the converse implication fails.

Restricting attention to continuous random variables, Schweizer and Wolff (see [9]) defined several nonparametric measures of monotone

dependence using copulas. By Sklar's theorem (see [10]), for any  $X$  and  $Y$  with continuous distribution functions  $F_X$  and  $F_Y$ , respectively, and joint distribution function  $F_{X,Y}$ , there exists a unique copula  $C$  such that

$$(1.4) \quad F_{X,Y}(x, y) = C(F_X(x), F_Y(y)).$$

The copula fully characterizes the dependence between  $X$  and  $Y$ , and provides a convenient tool for measuring stochastic dependence. For example, since  $X$  and  $Y$  are independent if and only if their connecting copula is the product copula  $P(x, y) = xy$ , it has been argued in [9] that *any* suitably normalized distance between  $C$  and  $P$ , in particular, any  $L^p$ -distance, should yield a symmetric nonparametric measure of dependence. Specifically, Schweizer and Wolff studied the  $L^1$ -distance given by

$$(1.5) \quad \sigma(X, Y) = 12 \int_{I^2} |C - P| d\lambda$$

where  $I = [0, 1]$  denotes the closed unit interval and  $\lambda$  the two-dimensional Lebesgue measure.

Strictly speaking, however,  $\sigma(X, Y)$  is a measure of *monotone* dependence because it attains its maximum of 1 if and only if  $X$  and  $Y$  are monotone dependent. This is easily seen since any copula lies (pointwise) between the lower and upper Fréchet-Hoeffding bounds, which are copulas themselves and correspond precisely to monotone decreasing and increasing dependence, respectively (see [9]). Moreover, if  $X$  and  $Y$  are m.c.d.,  $\sigma(X, Y)$  can attain any value in  $(0, 1]$ . This follows from the fact that the set of copulas linking m.c.d. random variables is dense in the set of all copulas with respect to the  $L^\infty$ -distance (see [6]), and, since copulas are continuous functions, with respect to any  $L^p$ -distance,  $p \geq 1$ . In fact, since copulas themselves are multivariate distribution functions with uniform margins, for which, due to uniform continuity, pointwise and uniform convergence coincide, this is equivalent to the above mentioned phenomenon that mutual complete dependence is not preserved under convergence in law.

In other words, none of the  $L^p$ -distances is capable of detecting mutual complete dependence except for the two extreme cases of monotone dependence. Therefore, the choice of the metrical distance function

used in the construction of a measure of dependence is crucial for its resulting properties.

In this paper, we argue that the inconsistency between mutual complete dependence and the  $L^p$ -distance neither weakens the concept of mutual complete dependence as the opposite of independence, nor does it imply that a measure of dependence should be restricted to monotone dependence. It rather suggests that convergence in law, or, alternatively, uniform convergence of the corresponding copulas, is an inappropriate concept for the construction of measures of dependence.

Instead of the  $L^p$ -norm, we propose to measure the distance between two copulas by a modified Sobolev norm given by

$$(1.6) \quad \|C\| = \left( \int_{I^2} |\nabla C|^2 d\lambda \right)^{1/2}$$

where  $\nabla$  denotes the gradient of the copula. This norm derives from a scalar product which, among other things, allows a straightforward representation via the  $*$ -product for copulas, introduced by Darsow et al. in [1]. Furthermore, this Sobolev norm turns out extremely advantageous since the degree of dependence between two continuous random variables  $X$  and  $Y$ , and, in particular, mutual complete dependence, can be determined by analytical properties of their copula. It follows that, in contrast to the  $L^p$ -distance, with respect to the Sobolev norm, a sequence of copulas corresponding to mutual complete dependence can only converge to a copula which itself links m.c.d. variables. Thus, mutual complete dependence cannot approximate any other kind of stochastic dependence.

Using this Sobolev norm we define a new nonparametric measure of dependence for two continuous random variables  $X$  and  $Y$  with copula  $C$ , given by

$$(1.7) \quad \omega(X, Y) = \left( 3\|C\|^2 - 2 \right)^{1/2} = \sqrt{3}\|C - P\|,$$

which represents the (normalized) Sobolev distance between  $C$  and the independence copula  $P$ . We show that the quantity  $\omega(X, Y)$  satisfies the following:

- (i)  $\omega(X, Y)$  is defined for any  $X$  and  $Y$ .
- (ii)  $\omega(X, Y) = \omega(Y, X)$ .

- (iii)  $0 \leq \omega(X, Y) \leq 1$ .
- (iv)  $\omega(X, Y) = 0$  if, and only if,  $X$  and  $Y$  are independent.
- (v)  $\omega(X, Y) = 1$  if, and only if,  $X$  and  $Y$  are m.c.d.
- (vi)  $\omega(X, Y) \in [1/2, 1]$  if  $Y$  is completely dependent on  $X$  (or vice versa).
- (vii) If  $f$  and  $g$  are a.s. strictly monotone functions on  $\text{Range}(X)$  and  $\text{Range}(Y)$ , respectively, then  $\omega(f(X), g(Y)) = \omega(X, Y)$ .
- (viii) If  $\{(X_n, Y_n)\}$  is a sequence of pairs of continuous random variables with copulas  $C_n$ , and if  $\lim_{n \rightarrow \infty} \|C_n - C\| = 0$ , then  $\lim_{n \rightarrow \infty} \omega(X_n, Y_n) = \omega(X, Y)$ .
- (ix) If  $X$  and  $Y$  are jointly normal with correlation coefficient  $\rho$ , then  $\omega(X, Y)$  is a strictly increasing function of  $|\rho|$ .

These properties show that  $\omega(X, Y)$  is a nonparametric measure of dependence, with extremes precisely at independence and mutual complete dependence.

The paper is organized as follows. Section 2 sets up the notation and briefly reviews some fundamental properties of copulas. Section 3 introduces the Sobolev scalar product for copulas and its corresponding norm and distance. We show that the scalar product allows a representation via the  $*$ -product. In Section 4 we turn to the statistical interpretation of the Sobolev norm for copulas, which leads naturally to a new nonparametric measure of dependence for two continuous random variables. Examples and comparisons are presented in Section 5.

**2. Basic properties of copulas.** Let  $I = [0, 1]$  be the closed unit interval and  $I^2 = [0, 1] \times [0, 1]$  the closed unit square.

**DEFINITION 2.1.** *A two-dimensional copula (or briefly, a copula) is a function  $C : I^2 \rightarrow I$  satisfying the conditions:*

- (i)  $C(x, 0) = C(0, y) = 0$  for all  $x, y \in I$ .
- (ii)  $C(x, 1) = x$  and  $C(1, y) = y$  for all  $x, y \in I$ .
- (iii)  $C(x_2, y_2) - C(x_2, y_1) - C(x_1, y_2) + C(x_1, y_1) \geq 0$  for all rectangles  $[x_1, x_2] \times [y_1, y_2] \subset I^2$ .

Let  $\mathfrak{C}$  denote the set of all (two-dimensional) copulas. Denote by  $\partial_i C$  the partial derivative of  $C \in \mathfrak{C}$  with respect to the  $i$ -th variable. The conditions in Definition 2.1 imply the following key properties of copulas; for a proof see, e.g., [7].

- PROPOSITION 2.2. (i)  $C$  is increasing in each argument.  
(ii)  $C$  is Lipschitz (and hence uniformly) continuous.  
(iii) For  $i \in \{1, 2\}$ ,  $\partial_i C$  exists a.e. on  $I^2$  with  $0 \leq \partial_i C(x, y) \leq 1$ .  
(iv) The functions  $t \mapsto \partial_1 C(x, t)$  and  $t \mapsto \partial_2 C(t, y)$  are defined and increasing a.e. on  $I$ .

There are three distinguished copulas, namely

$$\begin{aligned} C^-(x, y) &= \max(x + y - 1, 0), \\ C^+(x, y) &= \min(x, y), \\ P(x, y) &= xy. \end{aligned}$$

$C^+$  and  $C^-$  are called the Fréchet-Hoeffding upper and lower bound, respectively, since for any copula  $C$  and any  $(x, y) \in I^2$  we have

$$(2.1) \quad C^-(x, y) \leq C(x, y) \leq C^+(x, y).$$

The set  $\mathfrak{C}$  can be equipped with the  $*$ -multiplication (see [1]), defined by

$$(2.2) \quad (A * B)(x, y) = \int_0^1 \partial_2 A(x, t) \partial_1 B(t, y) dt.$$

$C^+$  and  $P$  are the unit and null element respectively, i.e., for any copula  $C$  we have

$$(2.3) \quad C^+ * C = C * C^+ = C,$$

$$(2.4) \quad P * C = C * P = P.$$

Denote by  $C^\top$  the transposed copula of  $C$  given by

$$(2.5) \quad C^\top(x, y) = C(y, x).$$

$C$  is called symmetric if  $C = C^\top$ . It is easy to see that for any  $A, B \in \mathfrak{C}$

$$(2.6) \quad (A * B)^\top = B^\top * A^\top.$$

A copula  $C$  is left invertible if there is a copula  $A$ , called a left inverse, such that  $A * C = C^+$ . It is right invertible if there is a copula  $A$ , called a right inverse, such that  $C * A = C^+$ . Left and right inverses are unique and correspond to the transposed copula  $C^\top$ , see [1, Theorem 7.1]. A copula is invertible if it is both left and right invertible and, in this case,  $C^\top$  is called the inverse of  $C$ .

**3. The Sobolev scalar product for copulas.** We denote by  $\cdot$  the Euclidean scalar product, by  $\|\cdot\|$  the Euclidean norm on  $\mathbb{R}^2$ , and by  $\lambda$  the 2-dimensional Lebesgue measure.

It follows immediately from Proposition 2.2 (iii), and has been noticed in [2], that

$$\mathfrak{C} \subset W^{1,p}(I^2, \mathbb{R})$$

for every  $p \in [1, \infty]$  where  $W^{1,p}(I^2, \mathbb{R})$  is the standard Sobolev space. However, it has not been exploited in this context that  $W^{1,2}(I^2, \mathbb{R})$  is a Hilbert space with respect to the usual  $W^{1,2}$ -scalar product

$$(3.1) \quad \langle f, g \rangle_{W^{1,2}} = \int_{I^2} fg \, d\lambda + \int_{I^2} \nabla f \cdot \nabla g \, d\lambda$$

so that the set of copulas,  $\mathfrak{C}$ , comes equipped with a scalar product structure.

There is, however, an even simpler way to define a scalar product for copulas. Since copulas satisfy  $C(0, 0) = 0$  we can actually forgo the first term. Indeed,

$$(3.2) \quad \langle f, g \rangle = \int_{I^2} \nabla f \cdot \nabla g \, d\lambda$$

defines a scalar product on the subspace

$$W_0^{1,2}(I^2, \mathbb{R}) = \{f \in W^{1,2}(I^2, \mathbb{R}) \mid f(0, 0) = 0\}$$

which contains  $\mathfrak{C}$ . The restriction of  $\langle \cdot, \cdot \rangle$  to  $\mathfrak{C}$  will be called the *Sobolev scalar product* for copulas. As usual, we define the corresponding Sobolev norm on  $\mathfrak{C}$  by

$$(3.3) \quad \|C\| = \left( \int_{I^2} |\nabla C|^2 \, d\lambda \right)^{1/2}$$

and the Sobolev distance function on  $\mathfrak{C} \times \mathfrak{C}$  by

$$(3.4) \quad d(A, B) = \left( \int_{I^2} |\nabla A - \nabla B|^2 \, d\lambda \right)^{1/2}.$$

**REMARK 3.1.** (i) Although the scalar products  $\langle \cdot, \cdot \rangle$  and  $\langle \cdot, \cdot \rangle_{W^{1,2}}$  are different, the resulting norms are equivalent on  $W_0^{1,2}(I^2, \mathbb{R})$  (where the nontrivial estimate is implied by Poincaré's inequality).

- (ii)  $(\mathfrak{C}, d)$  is a complete metric space, and the  $*$ -product is (jointly) continuous with respect to  $\| \cdot \|$ ; see [2].
- (iii) Besides, the Sobolev norm  $\| \cdot \|$  is reminiscent of the classical energy functional, which is well known in PDE's and differential geometry. In fact, one might call

$$E(C) = \frac{1}{2} \|C\|^2 = \frac{1}{2} \int_{I^n} |\nabla C|^2 d\lambda$$

the energy of a copula  $C$ .

We have seen that the Sobolev scalar product for copulas appears very naturally from analytical point of view. However, it also allows a representation via the  $*$ -product, defined in (2.2).

**THEOREM 3.2.** *For all  $A, B \in \mathfrak{C}$  we have the identity*

$$\begin{aligned} \langle A, B \rangle &= \int_0^1 (A^\top * B + A * B^\top)(t, t) dt \\ &= \int_0^1 (A^\top * B + B * A^\top)(t, t) dt. \end{aligned}$$

**PROOF.** It follows from (2.5) that

$$(3.5) \quad \begin{aligned} \partial_1 A^\top(x, y) &= \partial_2 A(y, x) \\ \partial_2 A^\top(x, y) &= \partial_1 A(y, x) \end{aligned}$$

Using (2.2) and (3.5) we can write

$$\begin{aligned} \int_0^1 \int_0^1 \partial_1 A(x, y) \partial_1 B(x, y) dx dy &= \int_0^1 \left( \int_0^1 \partial_2 A^\top(y, x) \partial_1 B(x, y) dx \right) dy \\ &= \int_0^1 (A^\top * B)(y, y) dy \\ \int_0^1 \int_0^1 \partial_2 A(x, y) \partial_2 B(x, y) dx dy &= \int_0^1 \left( \int_0^1 \partial_2 A(x, y) \partial_1 B^\top(y, x) dy \right) dx \\ &= \int_0^1 (A * B^\top)(x, x) dx. \end{aligned}$$

Adding up both terms we obtain the first identity.

The second equation in Theorem 3.2 is equivalent to

$$\int_0^1 (A * B^\top)(t, t) dt = \int_0^1 (B * A^\top)(t, t) dt$$

which follows from  $(A * B^\top)(t, t) = (A * B^\top)^\top(t, t) = (B * A^\top)(t, t)$ , where we have used (2.6).  $\square$

The representation in Theorem 3.2 becomes particularly simple for symmetric copulas.

**COROLLARY 3.3.** *If  $A, B \in \mathfrak{C}$  are symmetric, then*

$$\langle A, B \rangle = 2 \int_0^1 (A * B)(t, t) dt.$$

Theorem 3.2 yields upper and lower bounds for the scalar product of two copulas. More precisely, we have the following result.

**THEOREM 3.4.** *Let  $A, B \in \mathfrak{C}$ . Then*

$$\frac{1}{2} \leq \langle A, B \rangle \leq 1,$$

where both bounds are sharp.

**PROOF.** Theorem 3.2, in connection with the bounds for copulas given in (2.1), implies that

$$2 \int_0^1 C^-(t, t) dt \leq \langle A, B \rangle \leq 2 \int_0^1 C^+(t, t) dt.$$

Simple calculations yield  $\int_0^1 C^-(t, t) dt = 1/4$  and  $\int_0^1 C^+(t, t) dt = 1/2$ .

Finally, one easily computes that

$$(3.6) \quad \begin{aligned} \langle C^-, C^- \rangle &= \langle C^+, C^+ \rangle = 1 \\ \langle C^-, C^+ \rangle &= \frac{1}{2}. \end{aligned}$$

This shows that the bounds in the statement are sharp, and the proof is complete.  $\square$

**REMARK 3.5.** The diameter of  $(\mathfrak{C}, d)$  is 1. To prove this, consider the identity

$$d(A, B)^2 = \|A - B\|^2 = \|A\|^2 + \|B\|^2 - 2 \langle A, B \rangle.$$

It follows from Theorem 3.4 and (3.6) that  $d(A, B) \leq 1 = d(C^-, C^+)$ , which proves our claim.

**THEOREM 3.6.** *For all  $C \in \mathfrak{C}$ , the following hold:*

- (i)  $\langle C, P \rangle = 2/3$ .
- (ii)  $\|C - P\|^2 = \|C\|^2 - 2/3$ .
- (iii)  $2/3 \leq \|C\|^2 \leq 1$ .

**PROOF.** For (i), we remark that  $P = P^\top$ , so Theorem 3.2 and (2.4) imply

$$\langle P, C \rangle = \int_0^1 (P * C + C * P)(t, t) dt = 2 \int_0^1 P(t, t) dt = \frac{2}{3}.$$

This, in turn, proves (ii) because

$$\|C - P\|^2 = \|C\|^2 - 2\langle C, P \rangle + \|P\|^2 = \|C\|^2 - \frac{2}{3}.$$

Finally, (iii) is a consequence of (ii) and Theorem 3.4.  $\square$

**4. A nonparametric measure of dependence.** We now turn to the statistical interpretation of the Sobolev norm for copulas and the construction of a new nonparametric measure of dependence.

For the remaining part of the paper, let  $X$  and  $Y$  be continuous, real-valued random variables (on the same probability space) with copula  $C$ .

**LEMMA 4.1** ([1]). *The following statements are equivalent:*

- (i)  $Y$  is completely dependent on  $X$ .
- (ii)  $C$  is left invertible.
- (iii)  $\partial_1 C \in \{0, 1\}$  a.e.

*Consequently, the following are also equivalent:*

- (i)  $X$  and  $Y$  are mutually completely dependent.
- (ii)  $C$  is invertible.
- (iii)  $\partial_1 C, \partial_2 C \in \{0, 1\}$  a.e.

**PROOF.** Darsow et al. prove in [1, Theorem 11.1] that  $Y$  is completely dependent on  $X$  if and only if  $C$  is left invertible. Moreover, [1, Theorem 7.1] states that  $C$  has a left inverse if and only if for each  $y \in I$  one has  $\partial_1 C(\cdot, y) \in \{0, 1\}$  a.e. Actually, the proof shows that this

is tantamount to assuming that  $\partial_1 C(x, y) \in \{0, 1\}$  a.e. This proves the first part.

Analogous statements hold for right invertible copulas, from which the second part of the lemma follows.  $\square$

The next theorem describes the main results of this paper.

**THEOREM 4.2.** *The Sobolev norm for copulas satisfies  $\|C\|^2 \in [2/3, 1]$  for all  $C \in \mathfrak{C}$ . Moreover, the following assertions hold:*

- (i)  $\|C\|^2 = 2/3$  if, and only if,  $X$  and  $Y$  are independent.
- (ii)  $\|C\|^2 \in [3/4, 1]$  if  $Y$  is completely dependent on  $X$  (or vice versa).
- (iii)  $\|C\|^2 = 1$  if, and only if,  $X$  and  $Y$  are m.c.d.

*In terms of algebraic properties of  $C$ , these properties read as follows:*

- (i)  $\|C\|^2 = 2/3$  if, and only if,  $C = P$ .
- (ii)  $\|C\|^2 \in [3/4, 1]$  if  $C$  is left (or right) invertible.
- (iii)  $\|C\|^2 = 1$  if, and only if,  $C$  is invertible.

This result, together with the identity  $\|C - P\|^2 = \|C\|^2 - 2/3$ , expresses the astonishing fact that the Sobolev norm itself measures stochastic dependence, with extremes exactly at independence and mutual complete dependence. In addition, the Sobolev norm is able to detect that two random variables are not completely dependent.

**PROOF.** The foremost statement is contained in Theorem 3.6(iii).

The first assertion is an immediate consequence of Theorem 3.6(ii). It follows from (3.3) that

$$(4.1) \quad \|C\|^2 = \int_0^1 \int_0^1 (\partial_1 C(x, y))^2 dx dy + \int_0^1 \int_0^1 (\partial_2 C(x, y))^2 dx dy.$$

If  $Y$  is completely dependent on  $X$  we know from Lemma 4.1 that  $(\partial_1 C)^2 = \partial_1 C$  a.e., so the first summand in (4.1) is equal to

$$\int_0^1 \int_0^1 \partial_1 C(x, y) dx dy = \int_0^1 y dy = \frac{1}{2}.$$

The proof of Theorem 3.2 shows that the second term in (4.1) satisfies

$$\int_0^1 \int_0^1 (\partial_2 C(x, y))^2 dx dy = \int_0^1 (C * C^\top)(t, t) dt$$

which, since (2.1) implies  $C * C^\top \geq C^-$ , can be estimated by

$$\int_0^1 \int_0^1 (\partial_2 C(x, y))^2 dx dy \geq \int_0^1 C^-(t, t) dt = \frac{1}{4}.$$

This proves the second statement.

Finally, in view of Proposition 2.2 (iii), we have  $(\partial_i C)^2 \leq \partial_i C$  with equality if, and only if,  $\partial_i C \in \{0, 1\}$ . Consequently, (4.1) implies that

$$\|C\|^2 \leq \int_0^1 \int_0^1 \partial_1 C(x, y) dx dy + \int_0^1 \int_0^1 \partial_2 C(x, y) dx dy = \frac{1}{2} + \frac{1}{2} = 1$$

with equality if and only if  $\partial_i C \in \{0, 1\}$  a.e. By Lemma 4.1, the latter is equivalent to  $X$  and  $Y$  being m.c.d.  $\square$

**COROLLARY 4.3.** *The following are equivalent:*

- (i)  $X$  and  $Y$  are m.c.d.
- (ii)  $\|C\| = 1$ .
- (iii)  $\partial_1 C, \partial_2 C \in \{0, 1\}$  a.e.
- (iv)  $C$  is invertible, i.e.,  $C * C^\top = C^\top * C = C^+$ .
- (v)  $\int_0^1 (C * C^\top + C^\top * C)(t, t) dt = 1$ .

**PROOF.** This follows immediately from Lemma 4.1, Theorem 4.2 and Theorem 3.2.  $\square$

**COROLLARY 4.4.** *Let  $\{(X_n, Y_n)\}$  be a sequence of pairs of continuous random variables with copulas  $C_n$ . Then the following assertions hold:*

- (i) *If, for almost all  $n$ ,  $X_n$  and  $Y_n$  are m.c.d. and  $\lim_{n \rightarrow \infty} \|C_n - C\| = 0$ , then  $X$  and  $Y$  are m.c.d.*
- (ii) *If, for almost all  $n$ ,  $X_n$  is completely dependent on  $Y_n$ , or  $Y_n$  on  $X_n$ , and  $\lim_{n \rightarrow \infty} \|C_n - C\| = 0$ , then  $X$  and  $Y$  are not independent, i.e.,  $C \neq P$ .*

**PROOF.** Part (i) is an immediate consequence of Theorem 4.2 since  $\|C_n\| = 1$  and  $\lim_{n \rightarrow \infty} \|C_n - C\| = 0$  implies  $\|C\| = 1$ . An analogous argument proves (ii).  $\square$

Corollary 4.4 emphasizes the advantage of the Sobolev distance over the  $L^p$ -distances, as mentioned in the Introduction. While, in the uniform sense, any copula, in particular, the independence copula  $P$ , can be approximated by copulas of m.c.d. random variables, the Sobolev convergence preserves the property of mutual complete dependence. Hence, with respect to the Sobolev distance, mutual complete dependence cannot approximate any other kind of stochastic dependence. In fact, independence cannot even be approximated by completely dependent random variables.

Therefore, measuring the distance between copulas with the Sobolev norm resolves the disturbing phenomenon observed in [4, 6].

These remarkable statistical properties of the Sobolev norm lead immediately to the following definition.

**DEFINITION 4.5.** *Given two continuous random variables  $X, Y$  with copula  $C$ , we define*

$$\omega(X, Y) = \left(3\|C\|^2 - 2\right)^{1/2}.$$

In view of Theorem 3.6, the quantity  $\omega(X, Y)$  represents a normalized Sobolev distance of  $C$  from the independence copula  $P$ :

$$(4.2) \quad \omega(X, Y) = \sqrt{3} \|C - P\| = \frac{\|C - P\|}{\|\widehat{C} - P\|},$$

where  $\widehat{C}$  is any copula of m.c.d. variables. The normalization guarantees that  $\omega(X, Y) \in [0, 1]$ . Definition 4.5, however, makes clear that the Sobolev norm of  $C$  itself serves as a measure of dependence.

For symmetric  $C$  we may use Corollary 3.3 to write

$$(4.3) \quad \omega(X, Y) = \left(6 \int_0^1 (C * C)(t, t) dt - 2\right)^{1/2}.$$

**THEOREM 4.6.** *The quantity  $\omega(X, Y)$  has the following properties:*

- (i)  $\omega(X, Y)$  is defined for any  $X$  and  $Y$ .
- (ii)  $\omega(X, Y) = \omega(Y, X)$ .
- (iii)  $0 \leq \omega(X, Y) \leq 1$ .
- (iv)  $\omega(X, Y) = 0$  if, and only if,  $X$  and  $Y$  are independent.
- (v)  $\omega(X, Y) = 1$  if, and only if,  $X$  and  $Y$  are m.c.d.

- (vi)  $\omega(X, Y) \in [1/2, 1]$  if  $Y$  is completely dependent on  $X$  (or vice versa).
- (vii) If  $f$  and  $g$  are a.s. strictly monotone functions on  $\text{Range}(X)$  and  $\text{Range}(Y)$ , respectively, then  $\omega(f(X), g(Y)) = \omega(X, Y)$ .
- (viii) If  $\{(X_n, Y_n)\}$  is a sequence of pairs of continuous random variables with copulas  $C_n$ , and if  $\lim_{n \rightarrow \infty} \|C_n - C\| = 0$ , then  $\lim_{n \rightarrow \infty} \omega(X_n, Y_n) = \omega(X, Y)$ .

PROOF. Everything is obvious by definition, or follows from Theorem 4.2, except for (vi). Here we distinguish four different cases. For the sake of clarity, let  $C_{X,Y}$  denote the copula of  $X$  and  $Y$ .

If both  $f$  and  $g$  are increasing it is well known [7, Theorem 2.4.3] that  $C_{f(X),g(Y)} = C_{X,Y}$  which implies  $\omega(f(X), g(Y)) = \sqrt{3} \|C_{f(X),g(Y)} - P\| = \sqrt{3} \|C_{X,Y} - P\| = \omega(X, Y)$ .

If  $f$  is increasing and  $g$  is decreasing then  $C_{f(X),g(Y)}(x, y) = C_{X,Y}(x, 1-y)$ ; see [7, Theorem 2.4.4]. Therefore  $(C_{f(X),g(Y)} - P)(x, y) = (P - C_{X,Y})(x, 1-y)$  which, by the transformation formula for the Lebesgue measure, again implies  $\omega(f(X), g(Y)) = \omega(X, Y)$ . If  $f$  is decreasing and  $g$  is increasing, the result follows from interchanging  $f$  and  $g$  in the previous case.

The case when  $f$  and  $g$  are both decreasing can be shown similarly.  $\square$

REMARK 4.7. If  $X$  and  $Y$  are jointly normal with correlation coefficient  $\rho$ , then  $\omega(X, Y)$  is a strictly increasing function of  $|\rho|$  whose graph is shown in Figure 1.

**5. Examples and comparisons.** We conclude the paper with some examples clarifying the relationship between the measure of dependence  $\omega(X, Y)$  and the quantity  $\sigma(X, Y)$ , as defined in (1.5). We assume that the reader is familiar with the concept of singular copulas; for details we refer to [7].

EXAMPLE 5.1. Let  $\theta \in [0, 1]$ , and consider the singular copula  $C$  whose support consists of two line segments in  $I^2$ , one joining  $(0, 0)$  and  $(\theta, 1)$ , and the other joining  $(\theta, 1)$  and  $(1, 0)$  (see [7, Example 3.3]).

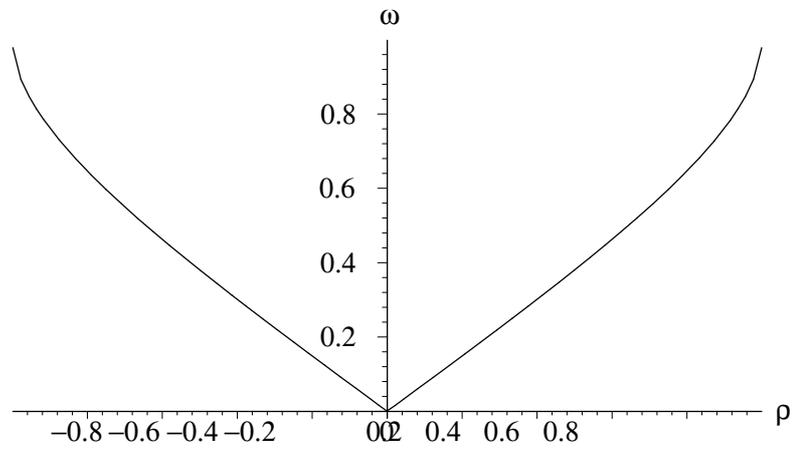


FIG 1.  $\omega(X, Y)$  as a function of  $\rho$  for jointly normal  $X, Y$

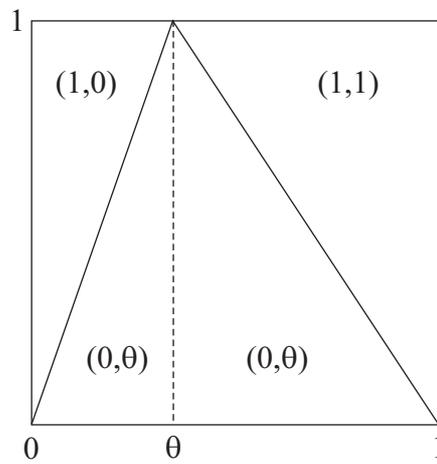


FIG 2. The gradient  $\nabla C$  of the copula  $C$  in Example 5.1

It follows that

$$C(x, y) = \begin{cases} x & \text{if } x \leq \theta y \\ \theta y & \text{if } \theta y < x < 1 - (1 - \theta)y \\ x + y - 1 & \text{if } 1 - (1 - \theta)y \leq x. \end{cases}$$

Clearly,  $Y$  is completely dependent on  $X$ , but not vice versa. Since probability mass  $\theta$  and  $1 - \theta$  is uniformly distributed on the first and second line segments, respectively, it is heuristically clear that the value  $\theta = 1/2$  describes the least dependent situation, whereas the limiting cases  $\theta = 0$  and  $\theta = 1$ , when  $C = C^-$  and  $C = C^+$ , respectively, correspond to mutual complete dependence.

This is perfectly reflected in the behavior of  $\omega(X, Y)$ . Indeed, a straightforward calculation (compare Fig. 2) shows that

$$\|C\|^2 = \frac{1}{2} \left( \theta - \frac{1}{2} \right)^2 + \frac{7}{8} \in \left[ \frac{7}{8}, 1 \right]$$

with the lowest and highest values attained precisely for  $\theta = 1/2$  and  $\theta \in \{0, 1\}$ , respectively. Consequently,  $\omega(X, Y)$  takes on its smallest value  $\sqrt{10}/4 \approx .79$  for  $\theta = 1/2$ .

The quantity  $\sigma(X, Y)$  shows the same qualitative behavior, however, its minimal value is .5.

**EXAMPLE 5.2.** Let  $\theta \in [0, 1]$ , and consider the singular copula  $C$  whose support consists of the two segments  $\{(x, 1 - x) \mid x \in [0, \theta] \cup [1 - \theta, 1]\}$  and the segment  $\{(x, x) \mid x \in [\theta, 1 - \theta]\}$  (see [7, Exercise 3.15]). It follows that

$$C(x, y) = \begin{cases} C^+(x, y) - \theta & \text{if } (x, y) \in [\theta, 1 - \theta]^2 \\ C^-(x, y) & \text{otherwise} \end{cases}$$

Now  $X$  and  $Y$  are mutually completely dependent so  $\omega(X, Y) = 1$ , regardless of the value of  $\theta$ .

In contrast,  $\sigma(X, Y)$  varies between 1 (for  $\theta \in \{0, 1\}$ ) and values around .46 (for  $\theta \approx .12$ ), indicating a definite degree of independence when, actually, there is none. Note that the copula from Example 5.1 with  $\theta = 1/2$  yields almost the same value for  $\sigma$ .

**Acknowledgements.** This work was partially supported by the German Science Foundation (Deutsche Forschungsgemeinschaft, Sonderforschungsbereich 475, Reduction of Complexity in Multivariate Data Structures). The second author gratefully acknowledges a scholarship from the Ruhr Graduate School in Economics and, in particular, from the Alfred Krupp von Bohlen and Halbach Foundation.

### References.

- [1] DARSOW, W., NGUYEN, B., AND OLSEN, E. (1992). Copulas and Markov processes. *Illinois J. Math.* **36**, 4, 600–642.
- [2] DARSOW, W. AND OLSEN, E. (1995). Norms for copulas. *Int. J. Math. and Math. Sci.* **18**, 3, 417–436.
- [3] GEBELEIN, H. (1941). Das statistische Problem der Korrelation als Variations- und Eigenwertproblem und sein Zusammenhang mit der Ausgleichsrechnung. *Z. Angew. Math. Mech.* **21**, 364–379.
- [4] KIMELDORF, G. AND SAMPSON, A. (1978). Monotone dependence. *Ann. Stat.* **6**, 4, 895–903.
- [5] LANCASTER, H. (1963). Correlation and complete dependence of random variables. *Ann. Math. Stat.* **34**, 1315–1321.
- [6] MIKUSIŃSKI, P., SHERWOOD, H., AND TAYLOR, M. (1992). Shuffles of min. *Stochastica* **13**, 1, 61–74.
- [7] NELSEN, R. (2006). *An Introduction to Copulas*, 2nd ed. John Wiley & Sons, New York.
- [8] RÉNYI, A. (1959). On measures of dependence. *Acta Math. Acad. Sci. Hungar.* **10**, 441–451.
- [9] SCHWEIZER, B. AND WOLFF, E. (1981). On nonparametric measures of dependence for random variables. *Ann. Stat.* **9**, 4, 879–885.
- [10] SKLAR, M. (1959). Fonctions de répartition à  $n$  dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris* **8**, 229–231.

FACHBEREICH MATHEMATIK  
UNIVERSITÄT DORTMUND  
VOGELPOTHSWEG 87  
44227 DORTMUND  
GERMANY  
E-MAIL: karl.f.siburg@math.uni-dortmund.de

FACHBEREICH STATISTIK  
UNIVERSITÄT DORTMUND  
VOGELPOTHSWEG 78  
44227 DORTMUND  
GERMANY  
E-MAIL: pavel.stoimenov@gmail.com