Multivariate comonotonicity

Giovanni Puccetti
Department of Mathematics for Decisions
University of Firenze
I–50134 Firenze, Italy,
giovanni.puccetti@unifi.it

Marco Scarsini
Dipartimento di Scienze Economiche e Aziendali
LUISS
Viale Romania 32
I–00197 Roma, Italy
and HEC, Paris
marco.scarsini@luiss.it

January 2009

Abstract

In this paper we consider several multivariate extensions of comonotonicity. We show that naive extensions do not enjoy some of the main properties of the univariate concept. In order to have these properties more structure is needed than in the univariate case.

AMS 2000 Subject Classification: Primary 60E05, Secondary 90B06.

Keywords: cyclical monotonicity, rearrangements, optimal coupling, copula, supermodular function.

1 Introduction

In recent years, the notion of comonotonicity has received a considerable attention in different fields. For instance, in the economics of insurance, the concept of comonotonic risks goes back to Borch (1962) and Arrow (1963, 1970, 1974), who, without using this term, proved that under some conditions an optimal insurance contract implies that the insurer and the insured have comonotonic wealths. In some models of modern decision theory the independence axiom of expected utility *à la* Savage has been replaced by a comonotonic independence axiom (see, e.g., Yaari (1987), Schmeidler (1989), Chew and Wakker (1996), Wakker (1996), and Kast and Lapied (2003)). In economic theory Landsberger and Meilijson (1994) show that for every allocation $(X_1,\ldots,X_n)$ of a random endowment $Y = \sum_{i=1}^{n} X_i$ among n agents, there is another comonotonic allocation $(X_1^*,\ldots,X_n^*)$ such that for every $i = 1,\ldots,n$, X_i^* dominates X_i in the sense of second degree stochastic dominance. Their result has been recently generalized by Ludkovski and Rüschendorf (2008). Carlier and Dana (2005) provide an existence theorem for a class of infinite-dimensional non-convex problems and sufficient conditions for monotonicity of optimal solutions. Among other things, in a model where agents have strictly Schur-concave utilities they prove that optimal contracts exist and agents' wealth are comonotonic. In actuarial sciences, the concept of univariate comonotonicity has several applications within the aggregation of insurance and financial risks, as discussed for instance in Dhaene et al. (2002a,b) and McNeil et al. (2005, Ch. 6). In the risk management of a portfolio $\boldsymbol{X} = (X_1, X_2)$ of losses with given marginal distributions, a financial institution is typically interested in the amount $\psi(\boldsymbol{X})$, representing the aggregate loss or the measure of the risk deriving from the portfolio. In particular, a regulator might require to calculate the worst-possible value attainable by $\psi(\boldsymbol{X})$. For many aggregating functionals ψ, this case is represented by comonotonicity among the risks.

In finance Galichon and Henry (2008a,b) propose a multivariate extension of coherent risk measures that involves a multivariate extension of the notion of comonotonicity, in the spirit of the present paper.

Consider the partially ordered space $\mathbb{R}\times\mathbb{R}$ with the component-wise order. A subset of $\mathbb{R}\times\mathbb{R}$ is called comonotonic if it is totally ordered. A random vector is called comonotonic if its support is a comonotonic set. It is well known that a random vector (X,Y) is comonotonic if and only if X and Y are nondecreasing functions of a common random factor, which can always be chosen to be the sum $X+Y$.

Given two univariate distribution functions F_X, F_Y, there always exists a comonotonic vector (X,Y) that has these marginal distributions. Its joint distribution is the upper Fréchet bound of the class of bivariate distributions with marginals F_X, F_Y (see Fréchet (1951)). The distribution of a comonotonic random vector with fixed marginals is unique. In a different language, the copula of a comonotonic random vector is the maximal copula (see, e.g., Sklar (1959), Schweizer and Sklar (1983), Nelsen (2006)).

An immediate consequence of the above properties is that, for any random variable X, the vector (X,X) is comonotonic.

It is important to notice that the definition of comonotonicity only relies on the total order structure of $\mathbb{R}$ and could be given for any random vectors with values in a product of totally ordered measurable spaces. Most of its properties would be valid even in this more general context.

The purpose of this paper is to study comonotonicity for pairs of random vectors, $(\boldsymbol{X}, \boldsymbol{Y})$. In this framework the multivariate marginal distributions of $\boldsymbol{X}$ and $\boldsymbol{Y}$ will be fixed and conditions for the existence of a comonotonic version will be studied. More formally, we want to study comonotonic vectors that take values in a product of partially ordered spaces.

The related problem of Fréchet bounds for multivariate marginals has been studied by Rüschendorf (1991a,b). Scarsini (1989) has studied copulae for measures on products of weakly ordered spaces. Furthermore Rüschendorf (2004, Section 5) has considered some cases of multivariate comonotonicity. Jouini and Napp (2003, 2004) have extended the concept of comonotonicity to dynamical settings.

We will consider different definitions of multivariate comonotonicity, trying to extend different features of the classical definition, and we will show that no definition satisfies all the properties of the original one. Some definitions do not guarantee the existence of a comonotonic random vector for any pair of multivariate marginals. Some other definitions do not guarantee uniqueness in distribution of the comonotonic random vector with fixed marginals. In order to guarantee these two properties more structure is necessary, hence a definition of multivariate comonotonicity (c-comonotonicity) that relies on the Hilbert-space nature of $\mathbb{R}^d$. This definition is based on the concept of *cyclical monotonicity*, introduced in Rockafellar (1970a). This concept of is equivalent to comonotonicity for $d = 1$.

Comonotonic random vectors are known to maximize a class of functionals. This idea goes back to the theory of rearrangements developed by Hardy et al. (1952), and has been extended in different contexts by different authors, e.g., Lorentz (1953), Cambanis et al. (1976), Meilijson and Nádas (1979), Tchen (1980), Rüschendorf (1980). C-comonotonic random vectors share similar properties (see, e.g., Brenier (1991)). Several other properties of cyclical comonotonicity are known in the literature (see, e.g., Cuesta-Albertos et al. (1993), Gangbo and McCann (1996), McCann (1995), Rüschendorf (2004)). We show that some of the good properties of c-comonotonicity are due to the stronger Hilbert-space structure its definition requires. A concept of multivariate comonotonicity, called μ-comonotonicity has been recently introduced by Galichon and Henry (2008a,b) with applications to measures of multivariate risk. This concept is variational in nature and will be compared to the other concepts of comonotonicity examined in the paper.

The paper is organized as follows. Section 2 considers the classical notion of comonotonicity for univariate marginals and its characterizations. In Section 3 several multivariate extensions are introduced and compared. Section 4 considers the maximization of some classes of functionals. Finally Section 5 consider two variational notions of multivariate comonotonicity.

2 Univariate marginals

In this section we review well-known results about the case of univariate marginals. We first fix some notation. Given two nonempty, partially ordered spaces $(\mathscr{X}, \leq_{\mathscr{X}})$ and $(\mathscr{Y}, \leq_{\mathscr{Y}})$, we will denote by $\lesssim$ the product partial order on $\mathscr{X} \times \mathscr{Y}$:

$$(x_1, y_1) \lesssim (x_2, y_2) \iff x_1 \leq_{\mathscr{X}} x_2 \text{ and } y_1 \leq_{\mathscr{Y}} y_2.$$

Except when explicitly said, we will consider the case $(\mathscr{X}, \leq_{\mathscr{X}}) = (\mathscr{Y}, \leq_{\mathscr{Y}}) = (\mathbb{R}^d, \leq)$, where $\leq$ is the natural component-wise order.

Given a nondecreasing function $\psi : \mathbb{R} \to \mathbb{R}$ its (right-continuous) generalized inverse is the function $\psi^{-1} : \mathbb{R} \to \overline{\mathbb{R}}$, defined as

$$\psi^{-1}(y) := \sup\{x \in \mathbb{R} : \psi(x) \leq y\}.$$

All random quantities will be defined on the probability space $(\Omega, \mathscr{A}, \mathbb{P})$. Given a random vector $\boldsymbol{X}$, $F_{\boldsymbol{X}}$ is its distribution function; $\boldsymbol{Y} \sim G$ means that G is the distribution function of $\boldsymbol{Y}$; $\boldsymbol{X} \overset{\text{dist}}{=} \boldsymbol{Y}$ means that $\boldsymbol{X}$ and $\boldsymbol{Y}$ have the same distribution; $\mathscr{U}[0,1]$ is the uniform distribution on the unit interval; finally $D := \{1, \dots, d\}$.

In the sequel we will often use the concept of copula and some of its basic properties. For this we refer the reader to Schweizer and Sklar (1983) and Nelsen (2006). We indicate by C_+ and C_- the upper and lower Fréchet bounds in the class of copulae:

$$\begin{aligned} C_+(u_1, \dots, u_d) &= \min(u_1, \dots, u_d), \\ C_-(u_1, \dots, u_d) &= \max\left(\sum_{i=1}^{d} u_i - d + 1, 0\right). \end{aligned}$$

Definition 2.1. The set $\Gamma \subset \mathbb{R} \times \mathbb{R}$ is said to be *comonotonic* if it is $\lesssim$-totally ordered, i.e. if for any $(x_1, y_1), (x_2, y_2) \in \Gamma$, either $(x_1, y_1) \lesssim (x_2, y_2)$, or $(x_1, y_1) \gtrsim (x_2, y_2)$.

Any random vector (X, Y) with comonotonic support is called comonotonic.

The following characterizations of comonotonic random vectors on $\mathbb{R} \times \mathbb{R}$ are well-known. They can be found for instance in Cuesta-Albertos et al. (1993, Proposition 2.1), Denneberg (1994, Proposition 4.5), Landsberger and Meilijson (1994, Section 2), and Dhaene et al. (2002b, Theorem 2).

Theorem 2.2. *The following statements are equivalent:*

(a) *the random vector* (X, Y) *is comonotonic;*

(b) $F_{(X,Y)}(x, y) = \min\{F_X(x), F_Y(y)\}$, *for all* $(x, y) \in \mathbb{R} \times \mathbb{R}$;

(c) $(X, Y) \overset{\text{dist}}{=} (F_X^{-1}(U), F_Y^{-1}(U))$, *where* $U \sim \mathscr{U}[0,1]$;

(d) *there exists a random variable* Z *and nondecreasing function* f_1, f_2 *such that* $(X, Y) \overset{\text{dist}}{=} (f_1(Z), f_2(Z))$;

(e) *X and Y are almost surely nondecreasing functions of $X+Y$;*

Note that all the characterizations of the theorem hold for an arbitrary (but finite) number of random variables.

Remark 2.3. In the case of univariate marginals the following properties hold:

(i) For every pair of marginals F_X, F_Y there exists a comonotonic random vector having these marginals.

(ii) The distribution of this comonotonic random vector is unique.

(iii) Only the total order structure of $\mathbb{R}$ is needed to define comonotonic random vectors. The definition could be given for random variables with values in any (measurable) totally ordered space.

(iv) For any random variable X, the vector (X, X) is comonotonic.

(v) If (X, Y) is comonotonic and $F_X = F_Y$, then $X = Y$ with probability one.

We now introduce the class of supermodular functions. The reader is referred to Topkis (1998) for properties of these functions.

Definition 2.4. A function $c : \mathbb{R}^n \to \mathbb{R}$ is said to be *supermodular* if

$$c(\boldsymbol{u} \wedge \boldsymbol{v}) + c(\boldsymbol{u} \vee \boldsymbol{v}) \geq c(\boldsymbol{u}) + c(\boldsymbol{v}), \text{ for all } \boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n, \tag{2.1}$$

where $\boldsymbol{u} \wedge \boldsymbol{v}$ is the component-wise minimum of $\boldsymbol{u}$ and $\boldsymbol{v}$, and $\boldsymbol{u} \vee \boldsymbol{v}$ is the component-wise maximum of $\boldsymbol{u}$ and $\boldsymbol{v}$. Call $\mathscr{S}_n$ the class of supermodular functions on $\mathbb{R}^n$.

When $n = 2$, a function $c : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is supermodular if and only if

$$c\left(x_1, y_1\right) + c\left(x_2, y_2\right) \geq c\left(x_1, y_2\right) + c\left(x_2, y_1\right), \text{ for all } x_2 \geq x_1,\ y_2 \geq y_1. \tag{2.2}$$

Comonotonic vectors maximize the expectation of supermodular functions over the class of all random vectors having the same marginals.

Theorem 2.5. *For every possible univariate distributions F_X and F_Y, denote by (X^*, Y^*) a comonotonic random vector such that $X^* \overset{\text{dist}}{=} F_X$ and $Y^* \overset{\text{dist}}{=} F_Y$, and let $c : \mathbb{R}^2 \to \mathbb{R}$ be right-continuous. Then*

$$\mathbb{E}\left[c\left(X^*, Y^*\right)\right] = \sup\left\{\mathbb{E}\left[c\left(\widetilde{X}, \widetilde{Y}\right)\right] : \widetilde{X} \sim F_X, \widetilde{Y} \sim F_Y\right\} \text{ for all } F_X \text{ and } F_Y, \tag{2.3}$$

if and only if the function $c \in \mathscr{S}_2$.

Proof. The *if* part follows from Rachev and Rüschendorf (1998, Remark 3.1.3), but many authors have derived the same result under different regularity conditions: see, e.g., Lorentz (1953), Cambanis et al. (1976, Theorem 1).

For the *only if* part, suppose that the right-continuous function c is not supermodular, i.e. it is possible to find $x_2 \geq x_1$ and $y_2 \geq y_1$ such that $c(x_1, y_1) + c(x_2, y_2) < c(x_1, y_2) + c(x_2, y_1)$. We will show that there exist two distribution functions F_X and F_Y such that a comonotonic vector having these marginals does not attain the supremum in (2.3). Let F_X assign mass 1/2 to the points x_1 and x_2, and F_Y assign mass 1/2 to the points y_1 and y_2. The random vector (X^*, Y^*) that assumes values (x_1, y_1) and (x_2, y_2) with probability 1/2 is comonotonic. Let now (X, Y) be a random vector that assumes values (x_1, y_2) and (x_2, y_1) with probability 1/2. Both vectors have the required marginals and

$$\mathbb{E}\left[c\left(X^*, Y^*\right)\right] = 1/2\left[c\left(x_1, y_1\right) + c\left(x_2, y_2\right)\right] < 1/2\left[c\left(x_1, y_2\right) + c\left(x_2, y_1\right)\right] = \mathbb{E}\left[c\left(X, Y\right)\right],$$

which contradicts (2.3). □

Since the function $c(x, y) = -(x-y)^2$ is supermodular, we obtain the well-known fact that comonotonic random vectors minimize the expected Euclidean distance among their components. Similar results go back to Dall'Aglio (1956) and are fundamental for the theory of probability metrics (see e.g., Zolotarev (1983) and Rachev (1991)).

3 Multivariate marginals

In this section we show different possible extensions of comonotonicity to (subsets of) the product space $\mathbb{R}^d \times \mathbb{R}^d$.

Unfortunately, we will see that trivial extensions of Definition 2.1 cannot guarantee at the same time existence and uniqueness of the law of a comonotonic vector having arbitrarily fixed multivariate marginal distributions.

3.1 s-comonotonicity

We start with the strongest definition of comonotonicity.

Definition 3.1. The set $\Gamma \subset \mathbb{R}^d \times \mathbb{R}^d$ is said to be *s(trongly)-comonotonic* if it is $\lesssim$-totally ordered, i.e. if for any $(\boldsymbol{x}_1, \boldsymbol{y}_1), (\boldsymbol{x}_2, \boldsymbol{y}_2) \in \Gamma$, either $(\boldsymbol{x}_1, \boldsymbol{y}_1) \lesssim (\boldsymbol{x}_2, \boldsymbol{y}_2)$, or $(\boldsymbol{x}_2, \boldsymbol{y}_2) \lesssim (\boldsymbol{x}_1, \boldsymbol{y}_1)$.

Any random vector $(\boldsymbol{X}, \boldsymbol{Y})$ with s-comonotonic support is called s-comonotonic.

For instance, for $d = 2$ consider the marginals $F_{\boldsymbol{X}} = C_+(F_{X_1}, F_{X_2})$ and $F_{\boldsymbol{Y}} = C_+(F_{Y_1}, F_{Y_2})$, for some univariate distributions F_{X_i}, F_{Y_i}, $i = 1, 2$. If $U \backsim \mathscr{U}[0,1]$, then the vector

$$\left(\left(F_{X_1}^{-1}(U), F_{X_2}^{-1}(U)\right), \left(F_{Y_1}^{-1}(U), F_{Y_2}^{-1}(U)\right)\right)$$

is s-comonotonic and has bivariate marginals $F_{\boldsymbol{X}}$ and $F_{\boldsymbol{Y}}$.

When $d = 1$, Definition 3.1 reduces to Definition 2.1.

Since the space $\mathbb{R}^d$ is not totally ordered when $d > 1$, s-comonotonicity imposes heavy constraints on the marginal distributions $F_{\boldsymbol{X}}$ and $F_{\boldsymbol{Y}}$.

Lemma 3.2. *If the set $\Gamma \subset \mathbb{R}^d \times \mathbb{R}^d$ is s-comonotonic, then the sets $\pi_i(\Gamma)$, $i = 1,2$ are $\leq$-totally ordered, π_1 and π_2 being the two natural projections from $\mathbb{R}^d \times \mathbb{R}^d$ to $\mathbb{R}^d$.*

Proof. Let $\boldsymbol{x}_1, \boldsymbol{x}_2$ be arbitrary vectors in $\pi_1(\Gamma)$. Then there exists $\boldsymbol{y}_1, \boldsymbol{y}_2 \in \mathbb{R}^d$ such that $(\boldsymbol{x}_1, \boldsymbol{y}_1), (\boldsymbol{x}_2, \boldsymbol{y}_2) \in \Gamma$. Since Γ is $\lesssim$-totally ordered, then either $(\boldsymbol{x}_1, \boldsymbol{y}_1) \lesssim (\boldsymbol{x}_2, \boldsymbol{y}_2)$, or $(\boldsymbol{x}_2, \boldsymbol{y}_2) \lesssim (\boldsymbol{x}_1, \boldsymbol{y}_1)$. In the first case we have that $\boldsymbol{x}_1 \leq \boldsymbol{x}_2$; in the second that $\boldsymbol{x}_2 \geq \boldsymbol{x}_1$. The $\leq$-total order of $\pi_2(\Gamma)$ is shown analogously. □

From Lemma 3.2 and Theorem 2.2(b), it follows that an s-comonotonic random vector $(\boldsymbol{X}, \boldsymbol{Y})$ has multivariate marginals of the form

$$F_{\boldsymbol{X}}(x_1, \ldots, x_d) = C_+\left(F_{X_1}(x_1), \ldots, F_{X_d}(x_d)\right), \tag{3.1a}$$
$$F_{\boldsymbol{Y}}(y_1, \ldots, y_d) = C_+\left(F_{Y_1}(y_1), \ldots, F_{Y_d}(y_d)\right), \tag{3.1b}$$

for some univariate distribution functions $F_{X_1}, \ldots, F_{X_d}$ and $F_{Y_1}, \ldots, F_{Y_d}$.

The following theorem characterizes s-comonotonicity and shows that $(\boldsymbol{X}, \boldsymbol{Y})$ is s-comonotonic if and only if the $2d$ random variables $X_1, \ldots, X_d, Y_1, \ldots, Y_d$ are all pairwise comonotonic in the sense of Definition 2.1.

Theorem 3.3. *Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be two random vectors with respective distributions $F_{\boldsymbol{X}}$ and $F_{\boldsymbol{Y}}$ of the form* (3.1). *The following statements are equivalent:*

(a) *the random vector $(\boldsymbol{X}, \boldsymbol{Y})$ is s-comonotonic;*

(b) $F_{(\boldsymbol{X},\boldsymbol{Y})}(x_1, \ldots, x_d, y_1, \ldots, y_d) = \min(F_{\boldsymbol{X}}(x_1, \ldots, x_d), F_{\boldsymbol{Y}}(y_1, \ldots, y_d))$, *for all* $((x_1, \ldots, x_d), (y_1, \ldots, y_d)) \in \mathbb{R}^d \times \mathbb{R}^d$;

(c) $(\boldsymbol{X}, \boldsymbol{Y}) \overset{\text{dist}}{=} \left(\left(F_{X_1}^{-1}(U), \ldots, F_{X_d}^{-1}(U)\right), \left(F_{Y_1}^{-1}(U), \ldots, F_{Y_d}^{-1}(U)\right)\right)$, *where* $U \sim \mathcal{U}[0,1]$;

(d) *there exist a random variable Z and nondecreasing functions $f_1, \ldots, f_d$, $g_1, \ldots, g_d$ such that* $(X, Y) \overset{\text{dist}}{=} \left((f_1(Z), \ldots, f_d(Z)), (g_1(Z), \ldots, g_d(Z))\right)$;

(e) *for all $i, j \in D$, X_i and Y_i are almost surely nondecreasing functions of $X_i + Y_j$ and $X_j + Y_i$, respectively;*

Proof. (a) ⇒ (b). Assume that the support Γ of $(\boldsymbol{X}, \boldsymbol{Y})$ is s-comonotonic and choose an arbitrary $(\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{R}^d \times \mathbb{R}^d$. Define the sets

$$A_1 := \{(\boldsymbol{u}, \boldsymbol{v}) \in \Gamma : \boldsymbol{u} \leq \boldsymbol{x}\} \quad \text{and} \quad A_2 := \{(\boldsymbol{u}, \boldsymbol{v}) \in \Gamma : \boldsymbol{v} \leq \boldsymbol{y}\}.$$

Note that $\mathbb{P}\left[\boldsymbol{X} \leq \boldsymbol{x}, \boldsymbol{Y} \leq \boldsymbol{y}\right] = \mathbb{P}[A_1 \cap A_2] \leq \min\{\mathbb{P}[A_1], \mathbb{P}[A_2]\}$.

We now prove that

$$\text{either } A_1 \subset A_2 \text{ or } A_2 \subset A_1. \tag{3.2}$$

Suppose on the contrary that $A_1 \not\subset A_2$ and $A_2 \not\subset A_1$. Then it is possible to find $(\boldsymbol{u}_1, \boldsymbol{v}_1) \in A_1 \setminus A_2$, and $(\boldsymbol{u}_2, \boldsymbol{v}_2) \in A_2 \setminus A_1$. By definition of A_1 and A_2 we have that $\boldsymbol{u}_1 \leq \boldsymbol{x}$ and $\boldsymbol{v}_2 \leq \boldsymbol{y}$.

Since both A_1 and A_2 are subsets of the $\precsim$-totally ordered Γ, either $(\boldsymbol{u}_1, \boldsymbol{v}_1) \precsim (\boldsymbol{u}_2, \boldsymbol{v}_2)$ or $(\boldsymbol{u}_2, \boldsymbol{v}_2) \precsim (\boldsymbol{u}_1, \boldsymbol{v}_1)$. The first alternative is not possible, since $\boldsymbol{v}_1 \le \boldsymbol{v}_2 \le \boldsymbol{y}$ would imply that $(\boldsymbol{u}_1, \boldsymbol{v}_1) \in A_2$. The second is not possible, either, since $\boldsymbol{u}_2 \le \boldsymbol{u}_1 \le \boldsymbol{x}$ would imply that $(\boldsymbol{u}_2, \boldsymbol{v}_2) \in A_1$. Therefore (3.2) holds.

By (3.2), $\mathbb{P}[A_1 \cap A_2] \ge \min\{\mathbb{P}[A_1], \mathbb{P}[A_2]\}$, hence $\mathbb{P}\left[\boldsymbol{X} \le \boldsymbol{x}, \boldsymbol{Y} \le \boldsymbol{y}\right] = \min\{\mathbb{P}[A_1], \mathbb{P}[A_2]\}$ from which (b) follows.

(b) ⇒ (c). Assume that $U \sim \mathscr{U}[0,1]$. Note that

$$
\begin{aligned}
&\mathbb{P}\left[\left(F_{X_1}^{-1}(U), \dots, F_{X_d}^{-1}(U)\right), \left(F_{Y_1}^{-1}(U), \dots, F_{Y_d}^{-1}(U)\right) \le (\boldsymbol{x}, \boldsymbol{y})\right] \\
&= \mathbb{P}\left[F_{X_1}^{-1}(U) \le x_1, \dots, F_{X_d}^{-1}(U) \le x_d, F_{Y_1}^{-1}(U) \le y_1, \dots, F_{Y_d}^{-1}(U) \le y_d\right] \\
&= \mathbb{P}\left[U \le F_{X_1}(x_1), \dots, U \le F_{X_d}(x_d), U \le F_{Y_1}(y_1), \dots, U \le F_{Y_d}(x_d)\right] \\
&= \mathbb{P}\left[U \le \min_{i \in D}\left(\min\left\{F_{X_i}(x_i), F_{Y_i}(y_i)\right\}\right)\right] \\
&= \min\left(\min_{i \in D}\left\{F_{X_i}(x_i)\right\}, \min_{i \in D}\left\{F_{Y_i}(y_i)\right\}\right) \\
&= \mathbb{P}\left[\boldsymbol{X} \le \boldsymbol{x}, \boldsymbol{Y} \le \boldsymbol{y}\right].
\end{aligned}
$$

(c) ⇒ (d). Straightforward.

(d)⇒ (e). If (d) is true, then, for all $i, j \in D$, the random vector (X_i, Y_j) is comonotonic. (e) then follows from Theorem (2.2)(e).

(e)⇒ (a). Suppose, on the contrary, that $(\boldsymbol{X}, \boldsymbol{Y})$ is not s-comonotonic. Recall that both $\boldsymbol{X}$ and $\boldsymbol{Y}$ have copula C_+, hence they have $\le$-totally ordered supports. Therefore, it is possible to find $(\boldsymbol{x}_1, \boldsymbol{y}_1), (\boldsymbol{x}_2, \boldsymbol{y}_2) \in \operatorname{supp}(\boldsymbol{X}, \boldsymbol{Y})$ such that $\boldsymbol{x}_1 \le \boldsymbol{x}_2$ and $\boldsymbol{y}_2 \ge \boldsymbol{y}_1$ with two strict inequalities, say, in the i-th and j-th coordinate, respectively. This implies that (X_i, Y_j) is not comonotonic and, by Theorem 2.2(e), this contradicts (e). □

Note that it is possible for a random vector $(\boldsymbol{X}, \boldsymbol{Y})$ to have a distribution as in Theorem 3.3 (b) without having marginals of the form (3.1). This happens in particular cases when the marginal distributions have big jumps (see Rüschendorf (2004) and references therein).

Definition 3.1 provides the simplest extension of comonotonicity to the product of multidimensional spaces, but the concept of dependence that it implies is extremely strong. Indeed, a reasonable requirement for any notion of comonotonicity is that the vector $(\boldsymbol{X}, \boldsymbol{X})$ be comonotonic for any choice of the d-variate distribution $F_{\boldsymbol{X}}$. This does not happen with s-comonotonicity, unless the copula of $F_{\boldsymbol{X}}$ is the upper Fréchet bound C_+. This is unsatisfactory, since we want to consider comonotonicity as a concept of dependence *between* two random vectors, and not *within* them.

Remark 3.4. The following properties hold:

(i) Given a pair of marginals $F_{\boldsymbol{X}}, F_{\boldsymbol{Y}}$ there exists an s-comonotonic random vector having these marginals if and only if both $F_{\boldsymbol{X}}$ and $F_{\boldsymbol{Y}}$ have copula C_+.

(ii) The distribution of this s-comonotonic random vector is unique.

(iii) Only the partial order structure of $\mathbb{R}^d$ is needed to define s-comonotonic random vectors. The definition could be given for random variables with values in any (measurable) partially ordered space.

(iv) Given a random vector $\boldsymbol{X}$, the vector $(\boldsymbol{X},\boldsymbol{X})$ is s-comonotonic only if $\boldsymbol{X}$ has copula C_+ (i.e., it is itself comonotonic).

(v) If $(\boldsymbol{X},\boldsymbol{Y})$ is s-comonotonic and $F_{\boldsymbol{X}} = F_{\boldsymbol{Y}}$, then $\boldsymbol{X} = \boldsymbol{Y}$ with probability one.

3.2 π-comonotonicity

We now consider a weaker concept of comonotonicity according to which the vector $(\boldsymbol{X},\boldsymbol{X})$ is always comonotonic. Let $A_i, B_i, i \in D$ be measurable subsets of the real line. Given a set $\Gamma \subset \left(\times_{i=1}^d A_i\right) \times \left(\times_{i=1}^d B_i\right)$, for all $i \in D$ we denote by $\Pi_i(\Gamma)$ its projection on the space $A_i \times B_i$.

Definition 3.5. The set $\Gamma \subset \left(\times_{i=1}^d A_i\right) \times \left(\times_{i=1}^d B_i\right)$ is said to be *π-comonotonic* if, for all $i \in D$, $\Pi_i(\Gamma)$ is comonotonic as a subset of $A_i \times B_i$. A random vector $(\boldsymbol{X},\boldsymbol{Y})$ with π-comonotonic support is called π-comonotonic.

When $d = 1$, Definition 3.5 is equivalent to Definitions 3.1 and 2.1. When $d > 1$, an s-comonotonic random vector is also π-comonotonic, but not vice versa. Rüschendorf (2004, Example 5.1) provides an example of a π-comonotonic random vector that is not s-comonotonic. We show a simpler version of this example. Let $F_{\boldsymbol{X}} = F_{\boldsymbol{Y}}(y_1, y_2) = C_-$, and let $U \sim \mathscr{U}[0,1]$. Then the random vector

$$((U, 1-U), (U, 1-U))$$

has bivariate marginals $F_{\boldsymbol{X}}$ and $F_{\boldsymbol{Y}}$, and is π-comonotonic, but not s-comonotonic.

Definition 3.5 imposes some constraints on the marginal distributions $F_{\boldsymbol{X}}$ and $F_{\boldsymbol{Y}}$ of a π-comonotonic random vector $(\boldsymbol{X},\boldsymbol{Y})$, but they are weaker than the ones imposed by s-comonotonicity.

Lemma 3.6. *If the random vector $(\boldsymbol{X},\boldsymbol{Y})$ is π-comonotonic, then its marginal distribution functions $F_{\boldsymbol{X}}$ and $F_{\boldsymbol{Y}}$ have a common copula.*

Proof. By Sklar's theorem for any random vector $(Z_1,\dots,Z_d)$ there exists a random vector $(U_1,\dots,U_d)$ such that, for all $i \in D$, $U_i \sim \mathscr{U}[0,1]$ and $(Z_1,\dots,Z_d) \stackrel{\text{dist}}{=} (F_{Z_1}^{-1}(U_1),\dots,F_{Z_d}^{-1}(U_d))$. Furthermore, if $(X_1,\dots,X_d) \stackrel{\text{dist}}{=} (F_{X_1}^{-1}(U_1),\dots,F_{X_d}^{-1}(U_d))$ and $(Y_1,\dots,Y_d) \stackrel{\text{dist}}{=} (F_{Y_1}^{-1}(U_1),\dots,F_{Y_d}^{-1}(U_d))$, then $F_{\boldsymbol{X}}$ and $F_{\boldsymbol{Y}}$ have a common copula.

Assume that the random vector $(\boldsymbol{X},\boldsymbol{Y}) = ((X_1,\dots,X_d),(Y_1,\dots,Y_d))$ is π-comonotonic. Then the random vector $(X_i, Y_i) := \Pi_i(\boldsymbol{X},\boldsymbol{Y})$ is comonotonic. By Theorem 2.2(c), for all $i \in D$, there exists $U_i \sim \mathscr{U}[0,1]$ such that $(X_i, Y_i) \sim \left(F_{X_i}^{-1}(U_i), F_{Y_i}^{-1}(U_i)\right)$. Therefore $(X_1,\dots,X_d) \stackrel{\text{dist}}{=} (F_{X_1}^{-1}(U_1),\dots,F_{X_d}^{-1}(U_d))$ and $(Y_1,\dots,Y_d) \stackrel{\text{dist}}{=} (F_{Y_1}^{-1}(U_1),\dots,F_{Y_d}^{-1}(U_d))$, hence $F_{\boldsymbol{X}}$ and $F_{\boldsymbol{Y}}$ have a common copula. □

Lemma 3.6 implies that a π-comonotonic vector $(\boldsymbol{X},\boldsymbol{Y})$ is forced to have marginals of the form

$$F_{\boldsymbol{X}}(x_1,\ldots,x_d) = C\left(F_{X_1}(x_1),\ldots,F_{X_d}(x_d)\right), \tag{3.3a}$$
$$F_{\boldsymbol{Y}}(y_1,\ldots,y_d) = C\left(F_{Y_1}(y_1),\ldots,F_{Y_d}(y_d)\right), \tag{3.3b}$$

for some univariate distribution functions $F_{X_1},\ldots,F_{X_d},F_{Y_1},\ldots,F_{Y_d}$, and a copula C. Note also that, if $\boldsymbol{X}$ and $\boldsymbol{Y}$ are s-comonotonic, then C in (3.3) is the upper Fréchet bound C_+.

The following theorem characterizes π-comonotonicity and shows that $(\boldsymbol{X},\boldsymbol{Y})$ is π-comonotonic if and only if $\boldsymbol{X}$ and $\boldsymbol{Y}$ have the same copula and every pair (X_i,Y_i) is comonotonic in the sense of Definition 2.1.

Theorem 3.7. *Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be two random vectors with respective distributions $F_{\boldsymbol{X}}$ and $F_{\boldsymbol{Y}}$ of the form* (3.3). *The following statements are equivalent:*

(a) *the random vector $(\boldsymbol{X},\boldsymbol{Y})$ is π-comonotonic;*

(b) $F_{(\boldsymbol{X},\boldsymbol{Y})}(\boldsymbol{x},\boldsymbol{y}) = C\left(\min\left\{F_{X_1}(x_1),F_{Y_1}(y_1)\right\},\ldots,\min\left\{F_{X_d}(x_d),F_{Y_d}(y_d)\right\}\right)$, *for all* $(\boldsymbol{x},\boldsymbol{y}) = \left((x_1,\ldots,x_d),(y_1,\ldots,y_d)\right)\in\mathbb{R}^d\times\mathbb{R}^d$*;*

(c) $(\boldsymbol{X},\boldsymbol{Y}) \overset{\text{dist}}{=} \left(\left(F^{-1}_{X_1}(U_1),\ldots,F^{-1}_{X_d}(U_d)\right),\left(F^{-1}_{Y_1}(U_1),\ldots,F^{-1}_{Y_d}(U_d)\right)\right)$, *where $\boldsymbol{U}=(U_1,\ldots,U_d)$ is a random vector having distribution C;*

(d) *there exists a random vector $\boldsymbol{Z}=(Z_1,\ldots,Z_d)$ and nondecreasing function $f_1,\ldots,f_d$, $g_1,\ldots,g_d$ such that $(\boldsymbol{X},\boldsymbol{Y}) \overset{\text{dist}}{=} \left((f_1(Z_1),\ldots,f_d(Z_d)),(g_1(Z_1),\ldots,g_d(Z_d))\right)$;*

(e) *for all $i\in D$, X_i and Y_i are almost surely nondecreasing functions of X_i+Y_i;*

Proof. (a) $\Rightarrow$ (b). Assume that the random vector $(\boldsymbol{X},\boldsymbol{Y}) = ((X_1,\ldots,X_d),(Y_1,\ldots,Y_d))$ is π-comonotonic. As noted in the proof of Lemma 3.6, there exists $U_i\sim\mathscr{U}[0,1]$ such that $(X_i,Y_i)\overset{\text{dist}}{=}\left(F^{-1}_{X_i}(U_i),F^{-1}_{Y_i}(U_i)\right)$; and this for all $i\in D$. Therefore we have that

$$\begin{aligned}
\mathbb{P}\left[\boldsymbol{X}\le\boldsymbol{x},\boldsymbol{Y}\le\boldsymbol{y}\right] &= \mathbb{P}\left[\times_{i=1}^d\left\{X_i\le x_i,Y_i\le y_i\right\}\right]\\
&= \mathbb{P}\left[\times_{i=1}^d\left\{F^{-1}_{X_i}(U_i)\le x_i,F^{-1}_{Y_i}(U_i)\le y_i\right\}\right]\\
&= \mathbb{P}\left[\times_{i=1}^d\left\{U_i\le F_{X_i}(x_i),U_i\le F_{Y_i}(y_i)\right\}\right]\\
&= \mathbb{P}\left[\times_{i=1}^d\left\{U_i\le \min\left\{F_{X_i}(x_i),F_{Y_i}(y_i)\right\}\right\}\right]\\
&= C\left(\min\left\{F_{X_1}(x_1),F_{Y_1}(y_1)\right\},\ldots,\min\left\{F_{X_d}(x_d),F_{Y_d}(y_d)\right\}\right).
\end{aligned}$$

(b) $\Rightarrow$ (c). Already noted in the proof of Lemma 3.6
(c) $\Rightarrow$ (d). Straightforward.
(d) $\Rightarrow$ (a) Assume (d). Then the support of $(\boldsymbol{X},\boldsymbol{Y})$ is the set

$$\left\{\left(\left(f_1(z_1),\ldots,f_d(z_d)\right),\left(g_1(z_1),\ldots,g_d(z_d)\right)\right),(z_1,\ldots,z_d)\in\operatorname{supp}(\boldsymbol{Z})\right\},$$

which is π-comonotonic.
(d) $\Longleftrightarrow$ (e) Note that (d) holds if and only if for all $i \in D$ the vector (X_i, Y_i) is comonotonic. The equivalence then follows from Theorem (2.2)(e). □

Corollary 3.8. *Let* (3.1) *hold. Then* $(\boldsymbol{X}, \boldsymbol{Y})$ *is* π*-comonotonic if and only if it is s-comonotonic.*

Even if weaker than s-comonotonicity, Definition 3.5 can be applied only to vectors having marginals with the same dependence structure. Therefore, we need to weaken the definition of comonotonicity even further.

Remark 3.9. The following properties hold:

(i) Given a pair of marginals $F_{\boldsymbol{X}}, F_{\boldsymbol{Y}}$ there exists a π-comonotonic random vector having these marginals if and only if both $F_{\boldsymbol{X}}$ and $F_{\boldsymbol{Y}}$ have the same copula.

(ii) The distribution of this π-comonotonic random vector is unique.

(iii) Only the fact that $\mathbb{R}^d$ is a product of totally ordered spaces is needed to define π-comonotonic random vectors. The definition could be given for random variables with values in any (measurable) product of totally ordered spaces.

(iv) Given a random vector $\boldsymbol{X}$, the vector $(\boldsymbol{X}, \boldsymbol{X})$ is π-comonotonic.

(v) If $(\boldsymbol{X}, \boldsymbol{Y})$ is π-comonotonic and $F_{\boldsymbol{X}} = F_{\boldsymbol{Y}}$, then $\boldsymbol{X} = \boldsymbol{Y}$ with probability one.

3.3 w-comonotonicity

In this section we show that any attempt at defining a multivariate concept of comonotonicity based on the component-wise ordering of random vectors leads to unsatisfactory results.

Definition 3.10. The set $\Gamma \subset \mathbb{R}^d \times \mathbb{R}^d$ is said to be *w(eakly)-comonotonic* if

$$\boldsymbol{x}_1 \leq \boldsymbol{x}_2 \iff \boldsymbol{y}_1 \leq \boldsymbol{y}_2, \text{ for any } \left(\boldsymbol{x}_1, \boldsymbol{y}_1\right), \left(\boldsymbol{x}_2, \boldsymbol{y}_2\right) \in \Gamma. \tag{3.4}$$

Any random vector $(\boldsymbol{X}, \boldsymbol{Y})$ with w-comonotonic support is called w-comonotonic.

The vector $(\boldsymbol{X}, \boldsymbol{Y})$ is w-comonotonic if and only if for every nondecreasing function $f : \mathbb{R}^d \to \mathbb{R}$ the vector $(f(\boldsymbol{X}), f(\boldsymbol{Y}))$ is comonotonic in the sense of Definition 2.1.

When $d = 1$, Definition 3.10 is equivalent to Definitions 3.5, 3.1, and 2.1. When $d > 1$, a π-comonotonic vector is also w-comonotonic, but not vice versa, as Example 3.11 shows.

Example 3.11. Assume that the distribution of $\boldsymbol{X} := (X_1, X_2)$ is a nonsymmetric copula C on $[0,1]^2$, and define the linear transformation $T : [0,1]^2 \to [0,1]^2$ as $T(x_1, x_2) := (x_2, x_1)$. By Definition 3.10 the vector $(\boldsymbol{X}, T(\boldsymbol{X}))$ is w-comonotonic. Denoting by C_T the distribution of $T(\boldsymbol{X})$, we have

$$\begin{aligned} C_T(u_1, u_2) &= \mathbb{P}\left[T_1(X_1, X_2) \leq u_1, T_2(X_1, X_2) \leq u_2\right] \\ &= \mathbb{P}\left[X_2 \leq u_1, X_1 \leq u_2\right] \\ &= C(u_2, u_1). \end{aligned}$$

Since C is nonsymmetric, $C \neq C_T$, hence $(\boldsymbol{X}, T(\boldsymbol{X}))$ is not π-comonotonic (see Figure 1).

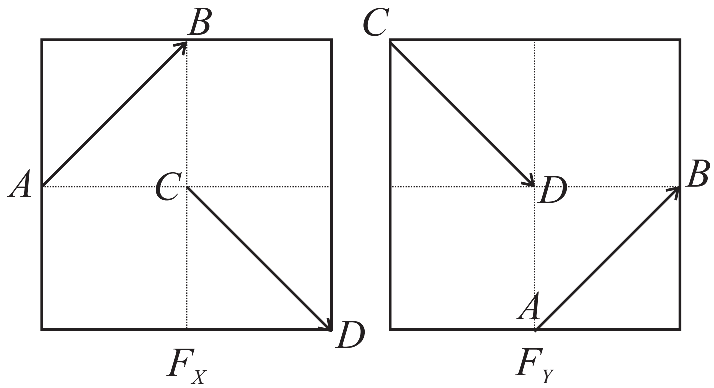

Figure 1: The support of a w-comonotonic vector which is not $\pi-$comonotonic. The segment AB on the left is mapped into the segment AB on the right. Similarly for segment CD.

Definition 3.10 includes the minimal intuitive requirement for a comonotonic random vector: high values for the first component go with high values for the other, whenever the two components are comparable. Nevertheless, this requirement still does not guarantee the existence of a w-comonotonic vector for an arbitrary choice of the multivariate marginals. For instance, if we assume that the first marginal $F_{\boldsymbol{X}}$ of a w-comonotonic vector $(\boldsymbol{X}, \boldsymbol{Y})$ has copula C_+, then the second marginal $F_{\boldsymbol{Y}}$ must have copula C_+, too.

Moreover, unlike the other definitions of s- and π-comonotonicity given above, Definition 3.10 does not assure uniqueness of the law of a w-comonotonic random vector having fixed multivariate marginals. The following example illustrates this crucial drawback.

Example 3.12. Let $d = 2$ and $F_X = F_Y = C_-$. The random vector $(\boldsymbol{X}, \boldsymbol{X})$ is w-comonotonic, but it is not the only one with F_X and F_Y as marginals. In fact, the vector $(\boldsymbol{X}, \mathbf{1} - \boldsymbol{X})$, where $\mathbf{1} := (1, 1)$, is w-comonotonic too (see Figure 2). The fact is a consequence of the weak constraint imposed by the definition. Since no pair of points in $\operatorname{supp}(C_-)$ is $\leq$-comparable, every random vector having these fixed marginals is w-comonotonic.

In fact, all our attempts to define a concept of comonotonicity based on the componentwise ordering of a support are doomed to fail. It is not possible to define a reasonable concept of comonotonicity based on (3.4) in any partially ordered space, as the following impossibility theorem shows.

Theorem 3.13. *Suppose that $(\mathscr{X}, \leq_{\mathscr{X}})$ and $(\mathscr{Y}, \leq_{\mathscr{Y}})$ are two partially ordered spaces containing at least two distinct points. If for any $\Upsilon_1 \subset \mathscr{X}, \Upsilon_2 \subset \mathscr{Y}$ it is possible to define a w-comonotonic set $\Gamma \subset \mathscr{X} \times \mathscr{Y}$ having Υ_1 and Υ_2 as its projections, then at least one of the following statement is true:*

(a) *$(\mathscr{X}, \leq_{\mathscr{X}})$ and $(\mathscr{Y}, \leq_{\mathscr{Y}})$ are totally ordered spaces;*

(b) *any set $\Gamma \subset \mathscr{X} \times \mathscr{Y}$ is w-comonotonic.*

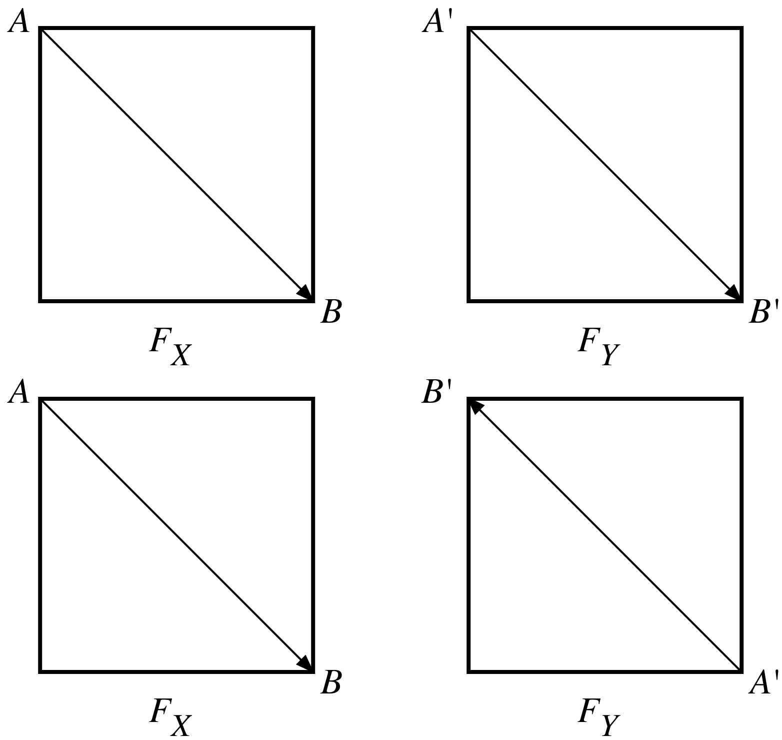

Figure 2: The support of two w-comonotonic vectors having the same marginals. In each case a point $\alpha A + (1 - \alpha B)$ on the left is mapped to the point $\alpha A' + (1 - \alpha B')$ on the right.

Proof. Assume that it is possible to find in one of the two spaces, say $\mathscr{X}$, two points $\boldsymbol{x}_1, \boldsymbol{x}_2$ with $\boldsymbol{x}_1 \leq_{\mathscr{X}} \boldsymbol{x}_2$ and in the other space $\mathscr{Y}$ two distinct points $\boldsymbol{y}_1, \boldsymbol{y}_2$ such that neither $\boldsymbol{y}_1 \leq \boldsymbol{y}_2$ nor $\boldsymbol{y}_2 \leq \boldsymbol{y}_1$ holds. Choose then $\Upsilon_1 := \{\boldsymbol{x}_1, \boldsymbol{x}_2\}$ and $\Upsilon_2 := \{\boldsymbol{y}_1, \boldsymbol{y}_2\}$. Note that Υ_1 is a totally ordered subset of $\mathscr{X}$, while no pairs of vectors are $\leq_{\mathscr{Y}}$-comparable in Υ_2. By definition of w-comotonicity, it is not possible to find a w-comonotonic set Γ in $\mathscr{X} \times \mathscr{Y}$ with projections Υ_1 and Υ_2.

Our initial assumption was then absurd, implying that either both spaces are totally ordered (hence (a) holds) or no pairs of vectors can be ordered in both of them. In this latter case, Definition 3.10 is always satisfied for any $\Upsilon_1 \subset \mathscr{X}$ and $\Upsilon_2 \subset \mathscr{Y}$ and then any set $\Gamma \in \mathscr{X} \times \mathscr{Y}$ is w-comonotonic, i.e., (b) holds. □

Translated in the language of probability, Theorem 3.13 states that every concept of comonotonicity including the requirement (3.4) on a partially ordered space, has to drop either the existence of a comonotonic vector for some choice of the marginals, or the uniqueness of its law.

A satisfactory concept of comonotonicity for univariate marginals is possible because $\mathbb{R}$ is totally ordered.

Remark 3.14. The following properties hold:

(i) Given a pair of marginals $F_{\boldsymbol{X}}, F_{\boldsymbol{Y}}$ the existence of a w-comonotonic random vector having these marginals is not always assured.

(ii) In general, the distribution of a w-comonotonic random vector with fixed marginals is not unique.

(iii) Only the partial order structure of $\mathbb{R}^d$ is needed to define w-comonotonic random vectors. The definition could be given for random variables with values in any (measurable) partially ordered space.

(iv) Given a random vector $\boldsymbol{X}$, the vector $(\boldsymbol{X}, \boldsymbol{X})$ is w-comonotonic.

(v) If $(\boldsymbol{X}, \boldsymbol{Y})$ is w-comonotonic and $F_{\boldsymbol{X}} = F_{\boldsymbol{Y}}$, then $\boldsymbol{X}$ is not necessarily equal to $\boldsymbol{Y}$ with probability one.

4 Maximization of functionals

We will now show that a result similar to Theorem 2.5 holds for s-comonotonic random vectors. To state the result we need to define a class of functions that includes the class of supermodular functions.

Definition 4.1. Call $\mathscr{S}_{\mathrm{s}}$ the class of functions $c : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ satisfying

$$c\left(\boldsymbol{x}_1, \boldsymbol{y}_1\right) + c\left(\boldsymbol{x}_2, \boldsymbol{y}_2\right) \geq c\left(\boldsymbol{x}_1, \boldsymbol{y}_2\right) + c\left(\boldsymbol{x}_2, \boldsymbol{y}_1\right) \tag{4.1}$$

for all $\boldsymbol{x}_2 \geq \boldsymbol{x}_1$ and $\boldsymbol{y}_2 \geq \boldsymbol{y}_1$.

Notice that for $d = 1$, the inequality in (4.1) reduces to (2.2), hence $\mathscr{S}_{\mathrm{s}} = \mathscr{S}_2$. The following lemma shows the relation between $\mathscr{S}_{2d}$ and $\mathscr{S}_{\mathrm{s}}$ for any d.

Lemma 4.2. $\mathscr{S}_{2d} \subset \mathscr{S}_{\mathrm{s}}$.

Proof. To prove that $\mathscr{S}_{2d} \subset \mathscr{S}_{\mathrm{s}}$ consider $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{y}_1, \boldsymbol{y}_2 \in \mathbb{R}^d$ such that $\boldsymbol{x}_2 \geq \boldsymbol{x}_1$ and $\boldsymbol{y}_2 \geq \boldsymbol{y}_1$. Call

$$\boldsymbol{u} = (\boldsymbol{x}_1, \boldsymbol{y}_2), \quad \boldsymbol{v} = (\boldsymbol{x}_2, \boldsymbol{y}_1).$$

Then

$$\boldsymbol{u} \wedge \boldsymbol{v} = \left(\boldsymbol{x}_1, \boldsymbol{y}_1\right), \quad \boldsymbol{u} \vee \boldsymbol{v} = \left(\boldsymbol{x}_2, \boldsymbol{y}_2\right).$$

Therefore (2.1) implies (4.1). □

The above inclusion is strict, as the following counterexample shows.

Consider the function $c : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ defined as

$$c(\boldsymbol{x}, \boldsymbol{y}) = \max(x_a, x_b) \max(y_a, y_b),$$

where $\boldsymbol{x} = (x_a, x_b)$ and $\boldsymbol{y} = (y_a, y_b)$.

Since the maximum is a monotone function, Hardy et al. (1952, Theorem 368) implies that, whenever $\boldsymbol{x}_2 \geq \boldsymbol{x}_1$ and $\boldsymbol{y}_2 \geq \boldsymbol{y}_1$, we have

$$\begin{aligned} c\left(\boldsymbol{x}_1, \boldsymbol{y}_1\right) + c\left(\boldsymbol{x}_2, \boldsymbol{y}_2\right) &= \max(x_{1a}, x_{1b}) \max(y_{1a}, y_{1b}) + \max(x_{2a}, x_{2b}) \max(y_{2a}, y_{2b}) \geq \\ &\max(x_{1a}, x_{1b}) \max(y_{2a}, y_{2b}) + \max(x_{2a}, x_{2b}) \max(y_{1a}, y_{1b}) = c\left(\boldsymbol{x}_1, \boldsymbol{y}_2\right) + c\left(\boldsymbol{x}_2, \boldsymbol{y}_1\right), \end{aligned} \tag{4.2}$$

with obvious meaning of the symbols. Hence the function c satisfies (4.1), i.e., $c \in \mathscr{S}_\mathrm{s}$.

To show that c is not supermodular consider the points $\boldsymbol{u} = (0,1,0,1)$, $\boldsymbol{v} = (1,0,1,0) \in \mathbb{R}^4$. We have $\boldsymbol{u} \wedge \boldsymbol{v} = (0,0,0,0)$ and $\boldsymbol{u} \vee \boldsymbol{v} = (1,1,1,1)$. Hence

$$c(\boldsymbol{u} \wedge \boldsymbol{v}) + c(\boldsymbol{u} \vee \boldsymbol{v}) = 0 + 1 < 1 + 1 = c(\boldsymbol{u}) + c(\boldsymbol{v}),$$

i.e., $c \notin \mathscr{S}_4$.

Theorem 4.3. *For every possible distributions $F_{\boldsymbol{X}}$ and $F_{\boldsymbol{Y}}$ of the form* (3.1), *let $\left(\boldsymbol{X}^*_\mathrm{s}, \boldsymbol{Y}^*_\mathrm{s}\right)$ be an s-comonotonic random vector having such marginal distributions, and let $c : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ be right-continuous. Then*

$$\mathbb{E}\left[c\left(\boldsymbol{X}^*_\mathrm{s}, \boldsymbol{Y}^*_\mathrm{s}\right)\right] = \sup\left\{\mathbb{E}\left[c\left(\widetilde{\boldsymbol{X}}, \widetilde{\boldsymbol{Y}}\right)\right] : \widetilde{\boldsymbol{X}} \sim F_{\boldsymbol{X}}, \widetilde{\boldsymbol{Y}} \sim F_{\boldsymbol{Y}}\right\} \text{ for all } F_{\boldsymbol{X}} \text{ and } F_{\boldsymbol{Y}}, \tag{4.3}$$

if and only if $c \in \mathscr{S}_\mathrm{s}$.

Proof. *If part*: Since $F_{\boldsymbol{X}}$ and $F_{\boldsymbol{Y}}$ are comonotonic distributions, by Theorem (2.2) (c), we can write

$$\begin{aligned} &\sup\left\{\mathbb{E}\left[c\left(\widetilde{\boldsymbol{X}}, \widetilde{\boldsymbol{Y}}\right)\right] : \widetilde{\boldsymbol{X}} \sim F_{\boldsymbol{X}}, \widetilde{\boldsymbol{Y}} \sim F_{\boldsymbol{Y}}\right\} \\ &= \sup\left\{\mathbb{E}\left[c\left(\left(F^{-1}_{X_1}(U_1), \dots, F^{-1}_{X_d}(U_1)\right), \left(F^{-1}_{Y_1}(U_2), \dots, F^{-1}_{Y_d}(U_2)\right)\right)\right] : U_1, U_2 \sim \mathscr{U}[0,1]\right\} \\ &= \sup\left\{\mathbb{E}\left[\hat{c}(U_1, U_2)\right] : U_1, U_2 \sim \mathscr{U}[0,1]\right\}, \end{aligned} \tag{4.4}$$

where $\hat{c} : [0,1] \times [0,1] \to \mathbb{R}$ is defined as

$$\hat{c}(z_1, z_2) := c\left(\left(F^{-1}_{X_1}(z_1), \dots, F^{-1}_{X_d}(z_1)\right), \left(F^{-1}_{Y_1}(z_2), \dots, F^{-1}_{Y_d}(z_2)\right)\right).$$

If $c \in \mathscr{S}_\mathrm{s}$ and is right-continuous, then it is easy to show that $\hat{c}$ is right continuous and supermodular. By Theorem 2.5, the supremum in (4.4) is reached when (U_1, U_2) is comonotonic, i.e. when $U_1 = U_2$ a.s., i.e., when the vector $\left(\widetilde{\boldsymbol{X}}, \widetilde{\boldsymbol{Y}}\right)$ is s-comonotonic.

Only if part: Suppose that $c \notin \mathscr{S}_\mathrm{s}$, i.e., that there exist $\boldsymbol{x}_2 \geq \boldsymbol{x}_1$ and $\boldsymbol{y}_2 \geq \boldsymbol{y}_1$ such that

$$c\left(\boldsymbol{x}_1, \boldsymbol{y}_1\right) + c\left(\boldsymbol{x}_2, \boldsymbol{y}_2\right) < c\left(\boldsymbol{x}_1, \boldsymbol{y}_2\right) + c\left(\boldsymbol{x}_2, \boldsymbol{y}_1\right).$$

We will show that there exist two distribution functions $F_{\boldsymbol{X}}$ and $F_{\boldsymbol{Y}}$ such that a comonotonic vector having these marginals does not attain the supremum in (4.3). Let $F_{\boldsymbol{X}}$ assign mass 1/2 to the points $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, and $F_{\boldsymbol{Y}}$ assign mass 1/2 to the points $\boldsymbol{y}_1$ and $\boldsymbol{y}_2$. The random vector $(\boldsymbol{X}^*, \boldsymbol{Y}^*)$ that assumes values $\left(\boldsymbol{x}_1, \boldsymbol{y}_1\right)$ and $\left(\boldsymbol{x}_2, \boldsymbol{y}_2\right)$ with probability 1/2 is s-comonotonic. Let now $(\boldsymbol{X}, \boldsymbol{Y})$ be a random vector that assumes values $\left(\boldsymbol{x}_1, \boldsymbol{y}_2\right)$ and $\left(\boldsymbol{x}_2, \boldsymbol{y}_1\right)$ with probability 1/2. Both vectors have the required marginals and

$$\mathbb{E}\left[c\left(\boldsymbol{X}^*, \boldsymbol{Y}^*\right)\right] = 1/2\left[c\left(\boldsymbol{x}_1, \boldsymbol{y}_1\right) + c\left(\boldsymbol{x}_2, \boldsymbol{y}_2\right)\right] < 1/2\left[c\left(\boldsymbol{x}_1, \boldsymbol{y}_2\right) + c\left(\boldsymbol{x}_2, \boldsymbol{y}_1\right)\right] = \mathbb{E}\left[c(\boldsymbol{X}, \boldsymbol{Y})\right],$$

which contradicts (4.3). □

We could not find an analogous result for π-comonotonic vectors. As the following counterexample shows, a π-comonotonic vector does not even maximize the expectation of a supermodular function.

Let $\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{X}^*, \boldsymbol{Y}^*$ have the same distribution, that assigns mass $1/2$ to the points $(0,1)$ and $(1,0)$.

Let now $(\boldsymbol{X}^*, \boldsymbol{Y}^*)$ take values $((0,1),(0,1))$ and $((1,0),(1,0))$ with probability $1/2$, and let $(\boldsymbol{X}, \boldsymbol{Y})$ take values $((0,1),(1,0))$ and $((1,0),(0,1))$ with probability $1/2$. The vector $(\boldsymbol{X}^*, \boldsymbol{Y}^*)$ is π-comonotonic. If

$$c(x_1, x_2, y_1, y_2) = x_2 y_1,$$

then c is supermodular since it is twice differentiable and all its mixed second derivatives are nonnegative. Nevertheless

$$\mathbb{E}[c(\boldsymbol{X}^*, \boldsymbol{Y}^*)] = 0 < 1 = \mathbb{E}[c(\boldsymbol{X}, \boldsymbol{Y})].$$

5 Variational multivariate comonotonicity

The concepts of s-comonotonicity and w-comonotonicity are based only on the partial order structure of $\mathbb{R}^d$, whereas π-comonotonicity involves also its product-space structure. In this section we will study different concepts of comonotonicity, that use the inner-product and are based on the maximization of some correlation.

5.1 c-comonotonicity

Definition 5.1. Given any two distributions $F_{\boldsymbol{X}}$ and $F_{\boldsymbol{Y}}$, the vector $\left(\boldsymbol{X}_c^*, \boldsymbol{Y}_c^*\right)$ is called *c(orrelation)-comonotonic* if

$$\mathbb{E}\left[\langle \boldsymbol{X}_c^*, \boldsymbol{Y}_c^* \rangle\right] = \sup\left\{\mathbb{E}\left[\langle \widetilde{\boldsymbol{X}}, \widetilde{\boldsymbol{Y}} \rangle\right] : \widetilde{\boldsymbol{X}} \sim F_{\boldsymbol{X}}, \widetilde{\boldsymbol{Y}} \sim F_{\boldsymbol{Y}}\right\}. \tag{5.1}$$

The concept of c-comonotonicity is strictly related to the concept of cyclical monotonicity studied by Rockafellar (1970a).

Definition 5.2. The set $\Gamma \subset \mathbb{R}^d \times \mathbb{R}^d$ is said to be *monotonic* if for any $(\boldsymbol{x}_0, \boldsymbol{y}_0), (\boldsymbol{x}_1, \boldsymbol{y}_1) \in \Gamma$,

$$\langle (\boldsymbol{x}_1 - \boldsymbol{x}_0), (\boldsymbol{y}_0 - \boldsymbol{y}_1) \rangle \le 0. \tag{5.2}$$

The set $\Gamma \subset \mathbb{R}^d \times \mathbb{R}^d$ is said to be *cyclically-monotonic* if for any $m > 1$ and $(\boldsymbol{x}_i, \boldsymbol{y}_i) \in \Gamma$, $i = 1, \ldots, m$, $\boldsymbol{x}_{m+1} := \boldsymbol{x}_1$,

$$\sum_{i=1}^{m} \langle (\boldsymbol{x}_{i+1} - \boldsymbol{x}_i), \boldsymbol{y}_i \rangle \le 0. \tag{5.3}$$

A multivalued mapping is called (cyclically) monotonic if its graph is (cyclically) monotonic. A *maximal* (cyclically) monotonic mapping is one whose graph is not properly contained in the graph of any other (cyclically) monotonic mapping.

A cyclically monotonic mapping is monotonic. The converse implication holds for $d = 1$. Monotonic operators have been studied by Zarantonello (1960, 1967), Minty (1962), and Brézis (1973).

Theorem 5.3. *A random vector $(\boldsymbol{X},\boldsymbol{Y})$ is c-comonotonic if and only if its support is cyclically monotonic.*

Theorem 5.3 has been proved by different authors, see, e.g., Rüschendorf and Rachev (1990, Theorem 1(b)) and Gangbo and McCann (1996, Corollary 2.4). A related, but different problem has been studied by Rüschendorf (1996), Beiglböck et al. (2008) and Schachermayer and Teichmann (2008). The new definition of c-comonotonicity includes the concepts of s- and π-comonotonicity as particular cases, whenever the marginals satisfy the appropriate constraints.

Lemma 5.4. *A π-comonotonic random vector is c-comonotonic.*

Proof. Assume that $(\boldsymbol{X},\boldsymbol{Y})$ is π-comonotonic. Choose an integer $m>2$ and arbitrary vectors $(\boldsymbol{x}_i,\boldsymbol{y}_i)\in\operatorname{supp}(\boldsymbol{X},\boldsymbol{Y})$, $i\in\{1,\dots,m\}$, $\boldsymbol{x}_{m+1}:=\boldsymbol{x}_1$ and denote by x^j the j-th component of the vector $\boldsymbol{x}$. We have that

$$\sum_{i=1}^{m}\langle(\boldsymbol{x}_{i+1}-\boldsymbol{x}_i),\boldsymbol{y}_i\rangle=\sum_{i=1}^{m}\sum_{j=1}^{d}\left(x_{i+1}^j-x_i^j\right)y_i^j=\sum_{j=1}^{d}\sum_{i=1}^{m}\left(x_{i+1}^j-x_i^j\right)y_i^j\le 0,$$

where the last inequality follows by noting that if $(\boldsymbol{X},\boldsymbol{Y})$ is π-comonotonic, then for all $j\in D$ the random vector (X_j,Y_j) is comonotonic, i.e. has a cyclically monotonic support in $\mathbb{R}\times\mathbb{R}$. □

A c-comonotonic vector which is neither $\pi-$comonotonic nor w-comonotonic can be found in Example 5.11 below. Moreover, the w-comonotonic vector $(\boldsymbol{X},\boldsymbol{1}-\boldsymbol{X})$ in Example 3.12 is not c-comonotonic. To prove this latter point, it is sufficient to verify that the vectors $(\boldsymbol{x}_1,\boldsymbol{y}_1)=((0,1),(1,0))$ and $(\boldsymbol{x}_2,\boldsymbol{y}_2)=((1,0),(0,1))$ in $\operatorname{supp}((\boldsymbol{X},\boldsymbol{1}-\boldsymbol{X}))$ do not satisfy condition (5.3) for $m=2$. This shows that the concept of c-comonotonicity is not based on the component-wise ordering of the coordinates in $\mathbb{R}^d$.

Contrary to the definitions of s-, π- and w-comonotonic vectors, Definition 5.1 does not impose any constraint on the marginal distributions of a c-comonotonic vector, as the following well known result shows.

Theorem 5.5. *Let $F_{\boldsymbol{X}}$ and $F_{\boldsymbol{Y}}$ be any two distributions on $\mathbb{R}^d$.*

(a) *There exists a c-comonotonic random vector $(\boldsymbol{X},\boldsymbol{Y})$ having marginals $F_{\boldsymbol{X}}$ and $F_{\boldsymbol{Y}}$.*

(b) *There exists a cyclically-monotonic set which includes all the supports of the c-comonotonic random vectors having marginals $F_{\boldsymbol{X}}$ and $F_{\boldsymbol{Y}}$.*

(c) *If the measure induced by either $F_{\boldsymbol{X}}$ or $F_{\boldsymbol{Y}}$ vanishes on all Borel subsets of Hausdorff dimension $d-1$, then all c-comonotonic random vectors have the same law.*

Theorem 5.5(a) and (c) are a suitable rewriting of McCann (1995, Theorem 6 and Corollary 14, respectively). This paper is mainly based on a fundamental result contained in Brenier (1991, Theorems 1.1 and 1.2). Point (b) is a consequence of Corollary 2.4 in Gangbo and McCann (1996) applied to the function $c(\boldsymbol{x},\boldsymbol{y})=\langle\boldsymbol{x},\boldsymbol{y}\rangle$. If the condition in (c) is violated, then several couplings can give rise to a c-comonotonic vector. For instance, let $U, V \sim \mathscr{U}[0,1]$, and consider the random vectors $\boldsymbol{X} = (U, 1-U)$ and $\boldsymbol{Y} = (V, V)$. It is not difficult to see that for every possible coupling the vector $(\boldsymbol{X}, \boldsymbol{Y})$ is always c-comonotonic. To prove this take $(\boldsymbol{x}_1, \boldsymbol{y}_1), \dots, (\boldsymbol{x}_m, \boldsymbol{y}_m)$ in the support of $(\boldsymbol{X}, \boldsymbol{Y})$, with $\boldsymbol{x}_i = (x_i, 1-x_i)$ and $\boldsymbol{y}_i = (y_i, y_i)$. Then

$$\sum_{i=1}^{m} \langle (\boldsymbol{x}_{i+1} - \boldsymbol{x}_i), \boldsymbol{y}_i \rangle = \sum_{i=1}^{m} \left[(x_{i+1} - x_i) y_i + (1 - x_{i+1} - 1 + x_i) y_i \right] = 0,$$

A different counterexample is provided by McCann (1995, Remark 5).

It is well known that the class of all $2d$-variate distribution functions having the fixed d-variate marginals $F_{\boldsymbol{X}}$ and $F_{\boldsymbol{Y}}$ contains the independence distribution (see Genest et al. (1995, Proposition A)). Theorem 5.5 states that this class always contains also a distribution that corresponds to a c-comonotonic vector.

A function $f : \mathbb{R}^d \to]-\infty, +\infty]$ is said to be lower semi-continuous if $\{\boldsymbol{x} \in \mathbb{R}^d : f(\boldsymbol{x}) \le t\}$ is closed in $\mathbb{R}^d$ for every $t \in \mathbb{R}$. Denote by $\mathscr{C}$ the class of lower semi-continuous, convex functions for which $\{\boldsymbol{x} \in \mathbb{R}^d : f(\boldsymbol{x}) < +\infty\} \ne \emptyset$. Given $f \in \mathscr{C}$, the *subdifferential* of f in $\boldsymbol{x}$ is the multivalued mapping defined by

$$\partial f(\boldsymbol{x}) = \left\{ \boldsymbol{y} \in \mathbb{R}^d : f(\boldsymbol{z}) - f(\boldsymbol{x}) \ge \langle (\boldsymbol{z} - \boldsymbol{x}), \boldsymbol{y} \rangle, \boldsymbol{z} \in \mathbb{R}^d \right\}.$$

Note that, when $\partial f(\boldsymbol{x})$ is a singleton, it reduces to the gradient of f, i.e. $\partial f(\boldsymbol{x}) = \{\nabla f(\boldsymbol{x})\}$.

For functions from $\mathbb{R}$ to $\mathbb{R}$ there exists a strict connection between convexity and monotonicity. A differentiable function is convex if and only if its derivative is nondecreasing. More generally a function is convex if its subdifferential is nondecreasing. The following result due to Rockafellar (1966, 1970b, Theorems A and B) provides a similar characterization for convex functions on $\mathbb{R}^d$.

Theorem 5.6. *If $f : \mathbb{R}^d \to \mathbb{R}^d$ is a lower semicontinuous convex function, then ∂f is a maximal monotonic operator from $\mathbb{R}^d$ to $\mathbb{R}^d$.*

Let $T : \mathbb{R}^d \to \mathbb{R}^d$ be a multivalued mapping. In order that there exist a lower semicontinuous convex function $f : \mathbb{R}^d \to \mathbb{R}^d$ such that $T = \partial f$, it is necessary and sufficient that T be a maximal cyclically monotonic operator.

As a consequence of Theorem 5.6 we obtain the following characterization of c-comonotonicity.

Theorem 5.7. *The following conditions are equivalent:*

(a) *The vector $(\boldsymbol{X}, \boldsymbol{Y})$ is c-comonotonic,*

(b) *$\boldsymbol{Y} \in \partial f(\boldsymbol{X})$ a.s. for some $f \in \mathscr{C}$,*

(c) *$\boldsymbol{X} \in \partial f(\boldsymbol{Y})$ a.s. for some $f \in \mathscr{C}$,*

Proof. Let $(\boldsymbol{X}, \boldsymbol{Y})$ be c-comonotonic. By Theorem 5.3 this happens if and only if its support Γ is cyclically monotonic. Using a result in Rockafellar (1966, page 501) or Rockafellar (1970a, page 27) we can actually assume that Γ is maximal cyclically monotonic. The multivalued mapping T defined as $T(y) = \{\boldsymbol{y} : (\boldsymbol{x}, \boldsymbol{y}) \in \Gamma\}$ is therefore maximally cyclically monotonic. The equivalence between (a) and (b) hence follows by Theorem 5.6 since for T to be maximally cyclically monotonic it is necessary and sufficient that there exist a lower semicontinuous function f such that $T \in \partial f(\boldsymbol{x})$. i.e. $Y \in \partial f(\boldsymbol{X})$ a.s.. In the previous proof, the random vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$ can be interchanged to prove the equivalence between (a) and (c). □

Theorem 5.7 shows a strong analogy with the univariate case. The next result and the counterexample that follows show that the analogy cannot be taken any further.

Theorem 5.8. *If the vector* $(\boldsymbol{X}, \boldsymbol{Y})$ *is c-comonotonic, then* $\boldsymbol{X} \in \partial f_1(\boldsymbol{Z})$ *a.s.,* $\boldsymbol{Y} \in \partial f_2(\boldsymbol{Z})$ *a.s. for some* $f_1, f_2 \in \mathscr{C}$ *and some random vector* $\boldsymbol{Z}$.

Proof. By Theorem 5.7 $(\boldsymbol{X}, \boldsymbol{Y})$ is c-comonotonic if and only if $\boldsymbol{X} \in \partial f(\boldsymbol{Y})$ a.s. for some $f \in \mathscr{C}$. Take now $\boldsymbol{Z} = \boldsymbol{Y}$ and $f_2(x) = \|x\|^2/2$. Hence $f_2 \in \mathscr{C}$ with $\partial f_2 = \{\mathrm{Id}\}$. □

Remark 5.9. Guillaume Carlier and an anonymous referee kindly pointed out to us that the converse of Theorem 5.8 does not hold. To show this, take $d = 2$, $\boldsymbol{Z} \sim \mathscr{N}(\boldsymbol{0}, I_2)$ a standard normal random vector. Define

$$\Sigma_1 = \begin{bmatrix} 1 & a \\ a & 2 \end{bmatrix}, \qquad \Sigma_2 = \begin{bmatrix} 1 & a \\ a & 1 \end{bmatrix},$$

with $0 < a < 1$. Since the matrices Σ_1 and Σ_2 are symmetric positive definite, we have that

$$f_1(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T \Sigma_1^{-1} \boldsymbol{x} \quad \text{and} \quad f_2(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T \Sigma_2 \boldsymbol{x}$$

are both differentiable convex functions with

$$\nabla f_1(\boldsymbol{x}) = \Sigma_1^{-1} \boldsymbol{x} \quad \text{and} \quad \nabla f_2(\boldsymbol{x}) = \Sigma_2 \boldsymbol{x}.$$

If we define

$$\boldsymbol{X} = \nabla f_1(\boldsymbol{Z}) \quad \text{and} \quad \boldsymbol{Y} = \nabla f_2(\boldsymbol{Z}),$$

then the thesis of Theorem 5.8 holds (since the functions f_1, f_2 are differentiable their subdifferential is restricted to their gradient). On the other hand $\boldsymbol{Z} = \Sigma_1 \boldsymbol{X}$, so $\boldsymbol{Y} = \Sigma_2 \Sigma_1 \boldsymbol{X}$, and

$$\Sigma_2 \Sigma_1 = \begin{bmatrix} 1 + a^2 & 3a \\ 2a & 2 + a^2 \end{bmatrix},$$

which is not symmetric, so $\Sigma_2 \Sigma_1 \boldsymbol{X}$ cannot be the gradient of a function. Therefore (b) cannot hold. This is a very striking departure from the univariate setting, where, by Theorem 2.2, comonotonicity of (X, Y) is equivalent to the fact that X and Y are nondecreasing functions of a common random variable Z.

When the function f is differentiable, Theorem 5.7 (b) gives the representation $(\boldsymbol{X}, T(\boldsymbol{X}))$ with $T = \nabla f$. This happens for instance when $F_{\boldsymbol{X}}$ vanishes on all Borel subsets of Hausdorff dimension $d - 1$; see McCann (1995).

Corollary 5.10. *The vector* $(\boldsymbol{X},\boldsymbol{X})$ *is* always *c-comonotonic.*

Proof. Choose $f(\boldsymbol{x}) = ||\boldsymbol{x}||^2/2$. Hence f is convex with $\partial f = \{\mathrm{Id}\}$, and the conditions of Theorem 5.7 (b) are satisfied. □

Cuesta-Albertos et al. (1993, Proposition 3.17) consider the following example of c-comonotonic vector, which, in our notation, is neither $\pi-$comonotonic nor w-comonotonic.

Example 5.11. Let $\boldsymbol{X} = (X_1, X_2)$ be a bivariate random vector having continuous marginal distributions F_{X_1}, F_{X_2} and copula $C(u_1,u_2) = u_1u_2$. The univariate marginals of $\boldsymbol{X}$ are therefore assumed to be independent. Now choose the operator $T : \mathbb{R}^2 \to \mathbb{R}^2$, defined as

$$T(x_1,x_2) = \left(\frac{e^{x_1}}{e^{x_1}+e^{x_2}}, \frac{e^{x_2}}{e^{x_1}+e^{x_2}}\right).$$

Since T is the gradient of the convex function $f : \mathbb{R}^2 \to \mathbb{R}, f(x_1,x_2) = \ln(e^{x_1}+e^{x_2})$, the vector $(\boldsymbol{X}, T(\boldsymbol{X}))$ is c-comonotonic. Note that the support of $T(\boldsymbol{X})$ is included in the decreasing curve $\{(t,1-t), t \in [0,1]\}$ in $\mathbb{R}^2$. By Theorem 2.5.5 in Nelsen (2006), the copula of $T(\boldsymbol{X})$ is C_- and therefore the vector $(\boldsymbol{X}, T(\boldsymbol{X}))$ cannot be $\pi-$comonotonic. Moreover, $(\boldsymbol{X}, T(\boldsymbol{X}))$ is not w-comonotonic, since no pair of vectors in $\mathrm{supp}(T(\boldsymbol{X}))$ is $\leq$-comparable while there exist such pairs in $\mathrm{supp}(\boldsymbol{X})$.

Theorem 5.7 gives an extension only of point (d) of Theorem 2.2. In general, it seems very difficult to find the law of a c-comonotonic vector (point (b) of Theorem 2.2), or the explicit form of the implied $d-$dimensional rearrangement (point (c)). The following theorem provides the extension of point (e) of Theorem 2.2.

Theorem 5.12. *If* $(\boldsymbol{X},\boldsymbol{Y})$ *is c-comonotonic, then both* $(\boldsymbol{X},\boldsymbol{X}+\boldsymbol{Y})$ *and* $(\boldsymbol{Y},\boldsymbol{X}+\boldsymbol{Y})$ *are c-comonotonic.*

Proof. We show that the support of $(\boldsymbol{X},\boldsymbol{X}+\boldsymbol{Y})$ is cyclically monotonic (the proof for $(\boldsymbol{Y},\boldsymbol{X}+\boldsymbol{Y})$ is analogous). Choose an integer $m > 2$ and arbitrary vectors $(\boldsymbol{x}_i, \boldsymbol{x}_i+\boldsymbol{y}_i) \in \mathrm{supp}(\boldsymbol{X},\boldsymbol{X}+\boldsymbol{Y}), i = 1,\dots,m, \boldsymbol{x}_{m+1} := \boldsymbol{x}_1$. Since the set $\mathrm{supp}(\boldsymbol{X},\boldsymbol{X})$ (see Remark 5.10) and $\mathrm{supp}(\boldsymbol{X},\boldsymbol{Y})$ (by assumption) are c-comonotonic, we have that

$$\sum_{i=1}^{m} \langle(\boldsymbol{x}_{i+1}-\boldsymbol{x}_i), \boldsymbol{x}_i+\boldsymbol{y}_i\rangle = \sum_{i=1}^{m} \langle(\boldsymbol{x}_{i+1}-\boldsymbol{x}_i), \boldsymbol{x}_i\rangle + \sum_{i=1}^{m} \langle(\boldsymbol{x}_{i+1}-\boldsymbol{x}_i), \boldsymbol{y}_i\rangle \leq 0,$$

i.e. $(\boldsymbol{X},\boldsymbol{X}+\boldsymbol{Y})$ is c-comonotonic. □

Using the concept of cyclical monotonicity, it is easy to show that the converse of Theorem 5.12 holds when $d = 1$, coherently with point (e) in Theorem 2.2. Unfortunately, this is no more true when $d > 1$, as the next counterexample shows.

Example 5.13. Let $(\boldsymbol{X},\boldsymbol{Y})$ in $\mathbb{R}^2 \times \mathbb{R}^2$ be the random vector uniformly distributed on the two points $(\boldsymbol{x}_1,\boldsymbol{y}_1) = ((0,1),(2,1))$ and $(\boldsymbol{x}_2,\boldsymbol{y}_2) = ((1,0),(0,0))$. The vector $(\boldsymbol{X}+\boldsymbol{Y})$ on $\mathbb{R}^2$

is then uniformly distributed on the two points $\boldsymbol{z}_1 = (2,2)$ and $\boldsymbol{z}_2 = (1,0)$. To prove that $(\boldsymbol{X}, \boldsymbol{X}+\boldsymbol{Y})$ is c-comonotonic, it is sufficient to note that

$$\langle(\boldsymbol{x}_2 - \boldsymbol{x}_1), \boldsymbol{z}_1\rangle + \langle(\boldsymbol{x}_1 - \boldsymbol{x}_2), \boldsymbol{z}_2\rangle = \langle(1,-1),(2,2)\rangle + \langle(-1,1),(1,0)\rangle = -1 \le 0.$$

Analogously, we have

$$\langle(\boldsymbol{y}_2 - \boldsymbol{y}_1), \boldsymbol{z}_1\rangle + \langle(\boldsymbol{y}_1 - \boldsymbol{y}_2), \boldsymbol{z}_2\rangle = \langle(-2,1),(2,2)\rangle + \langle(2,1),(1,0)\rangle = -4 \le 0,$$

hence also $(\boldsymbol{Y}, \boldsymbol{X}+\boldsymbol{Y})$ is c-comonotonic. For $(\boldsymbol{X}, \boldsymbol{Y})$, we find

$$\langle(\boldsymbol{x}_2 - \boldsymbol{x}_1), \boldsymbol{y}_1\rangle + \langle(\boldsymbol{x}_1 - \boldsymbol{x}_2), \boldsymbol{y}_2\rangle = \langle(1,-1),(2,1)\rangle + \langle(-1,1),(0,0)\rangle = 1 > 0,$$

hence $(\boldsymbol{X}, \boldsymbol{Y})$ is *not* c-comonotonic.

Even if in general c-comonotonic vectors are not unique, in the special case of equal marginals uniqueness holds.

Proposition 5.14. *If $(\boldsymbol{X}, \boldsymbol{Y})$ is c-comonotonic and $F_{\boldsymbol{X}} = F_{\boldsymbol{Y}}$, then $\boldsymbol{X} = \boldsymbol{Y}$ with probability one.*

Proof. Let $F_{\boldsymbol{X}} = F_{\boldsymbol{Y}}$, and let $(\boldsymbol{X}, \boldsymbol{Y})$ be c-comonotonic with $\boldsymbol{X} \neq \boldsymbol{Y}$. Then for all integer $m > 1$ and all $(\boldsymbol{x}_i, \boldsymbol{y}_i)$, $i = 1, \dots, m$, $\boldsymbol{x}_{m+1} := \boldsymbol{x}_1$, in the support of $(\boldsymbol{X}, \boldsymbol{Y})$ (5.3) holds.

Since $F_{\boldsymbol{X}} = F_{\boldsymbol{Y}}$, we can always choose the x_i so that $\boldsymbol{y}_i = \boldsymbol{x}_{i+1}$. Therefore, given (5.3), we have

$$\begin{aligned}
0 &\ge \sum_{i=1}^{m} \langle(\boldsymbol{x}_{i+1} - \boldsymbol{x}_i), \boldsymbol{y}_i\rangle \\
&= \sum_{i=1}^{m} \langle(\boldsymbol{x}_{i+1} - \boldsymbol{x}_i), \boldsymbol{x}_{i+1}\rangle \\
&= \sum_{i=1}^{m} \sum_{j=1}^{d} \left(x_{i+1}^j - x_i^j\right) x_{i+1}^j \\
&= \sum_{j=1}^{d} \sum_{i=1}^{m} \left(x_{i+1}^j - x_i^j\right) x_{i+1}^j \\
&= \sum_{j=1}^{d} \left[\sum_{i=1}^{m} x_{i+1}^j x_{i+1}^j - \sum_{i=1}^{m} x_{i+1}^j x_i^j\right].
\end{aligned}$$

By Hardy et al. (1952, Theorem 368), the expression in square brackets is nonnegative, hence for all $j = 1, \dots, d$ we have

$$\sum_{i=1}^{m} x_{i+1}^j x_{i+1}^j = \sum_{i=1}^{m} x_{i+1}^j x_i^j.$$

This is possible if and only if $(x_1^j, \dots, x_m^j)$ and $(x_2^j, \dots, x_{m+1}^j)$ are equally arranged. Since $x_{m+1}^j = x_1^j$, this is a contradiction. □

Remark 5.15. The following properties hold:

(i) Given a pair of marginals $F_{\boldsymbol{X}}, F_{\boldsymbol{Y}}$ there exists a c-comonotonic random vector having these marginals.

(ii) The distribution of a c-comonotonic random vector with fixed marginals is unique if at least one of the marginals is continuous.

(iii) The concept of inner product is necessary to define c-comonotonic random vectors. The definition could be given for random variables with values in any Hilbert space.

(iv) Given a random vector $\boldsymbol{X}$, the vector $(\boldsymbol{X}, \boldsymbol{X})$ is c-comonotonic.

(v) If $(\boldsymbol{X}, \boldsymbol{Y})$ is c-comonotonic and $F_{\boldsymbol{X}} = F_{\boldsymbol{Y}}$, then $\boldsymbol{X} = \boldsymbol{Y}$ with probability one.

5.2 μ-comonotonicity

Galichon and Henry (2008a,b) have recently proposed a variational concept of multivariate comonotonicity, called μ-comonotonicity. Related ideas will be considered in Section 4.

In this subsection, for the sake of simplicity, we identify a probability measure on $\left(\mathbb{R}^d, \mathrm{Bor}(\mathbb{R}^d)\right)$ with its distribution function.

Definition 5.16. Let μ be a probability measure on $\mathbb{R}^d$ that vanishes on Borel subsets of Hausdorff dimension $d-1$. The vector $(\boldsymbol{X}, \boldsymbol{Y}) \in L^\infty$ is called μ-comonotonic if for some random vector $\boldsymbol{V}$ distributed according to μ we have

$$\boldsymbol{V} \in \arg\max_{\tilde{\boldsymbol{V}}} \left\{\mathbb{E}[\langle \boldsymbol{X}, \tilde{\boldsymbol{V}}\rangle], \tilde{\boldsymbol{V}} \sim \mu\right\},$$
$$\boldsymbol{V} \in \arg\max_{\tilde{\boldsymbol{V}}} \left\{\mathbb{E}[\langle \boldsymbol{Y}, \tilde{\boldsymbol{V}}\rangle], \tilde{\boldsymbol{V}} \sim \mu\right\}.$$

The vector $(\boldsymbol{X}, \boldsymbol{Y})$ is μ-comonotonic if and only if there exists a random vector $\boldsymbol{Z} \sim \mu$ and two lower semi-continuous convex functions f and g such that $\boldsymbol{X} = \nabla f(\boldsymbol{Z})$ and $\boldsymbol{Y} = \nabla g(\boldsymbol{Z})$ almost surely. This idea is related to the concept of Pseudo-Wasserstein distance induced by μ, elaborated by Ambrosio et al. (2008, Section 3).

This definition generalizes the univariate definition of comonotonicity since when $d = 1$ a vector (X, Y) is comonotonic if and only if there exists a random variable Z^* having a nonatomic distribution μ such that

$$Z^* \in \arg\max_{Z} \left\{\mathbb{E}[ZX], Z \sim \mu\right\},$$
$$Z^* \in \arg\max_{Z} \left\{\mathbb{E}[ZY], Z \sim \mu\right\},$$

which happens if and only if there exists a random variable $U^* \sim \mathscr{U}[0,1]$ such that

$$U^* \in \arg\max_{U} \{\mathbb{E}[UX], U \sim \mathscr{U}[0,1]\},$$
$$U^* \in \arg\max_{U} \{\mathbb{E}[UY], U \sim \mathscr{U}[0,1]\}.$$

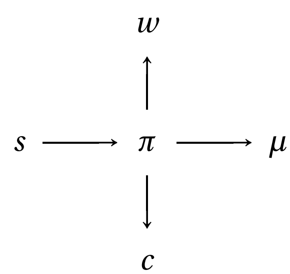

Table 5.1: Relationships between multivariate comonotonic concepts. Converse implications do not hold.

The main difference between the univariate and the multivariate case is that for $d = 1$ the choice of μ is irrelevant (for instance it can always be chosen to be uniform), for $d > 1$ it is not.

Theorem 5.17. *Let $(\boldsymbol{X}, \boldsymbol{Y})$ be π-comonotonic, and let the marginal distributions $F_{\boldsymbol{X}}$ and $F_{\boldsymbol{Y}}$ be of the form* (3.3). *Then $(\boldsymbol{X}, \boldsymbol{Y})$ is μ-comonotonic for $\mu = C$.*

Proof. Let $\boldsymbol{U} \sim C$. If we define for $i \in D$

$$X_i = F_{X_i}^{-1}(U_i), \quad Y_i = F_{Y_i}^{-1}(U_i)$$

by Theorem 3.7 the vector $(\boldsymbol{X}, \boldsymbol{Y})$ is π-comonotonic and has marginals $F_{\boldsymbol{X}}$ and $F_{\boldsymbol{Y}}$.

Furthermore, for $i \in D$

$$U_i \in \arg\max_{V} \{\mathbb{E}[VX], V \sim \mathscr{U}[0,1]\},$$
$$U_i \in \arg\max_{V} \{\mathbb{E}[VY], V \sim \mathscr{U}[0,1]\}.$$

By summing over $i \in D$ we obtain

$$\boldsymbol{U} \in \arg\max_{\boldsymbol{V}} \{\mathbb{E}[\langle \boldsymbol{X}, \boldsymbol{V} \rangle], \boldsymbol{V} \sim C\},$$
$$\boldsymbol{U} \in \arg\max_{\boldsymbol{V}} \{\mathbb{E}[\langle \boldsymbol{Y}, \boldsymbol{V} \rangle], \boldsymbol{V} \sim C\},$$

i.e., $(\boldsymbol{X}, \boldsymbol{Y})$ is μ-comonotonic with $\mu = C$. □

TABLE 5.1 ABOUT HERE

Acknowledgments

We are greatly indebted to Guillaume Carlier for pointing out a flaw in a previous version of this paper, and to an anonymous referee for a very accurate reading and a host of suggestions (including one concerning the above-mentioned flaw).

References

AMBROSIO, L., GIGLI, N., and SAVARÉ, G. (2008) *Gradient Glows in Metric Spaces and in the Space of Probability Measures.* Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition.

ARROW, K. J. (1963) Uncertainty and the welfare of medical care. *Amer. Econom. Rev.* **53**, 941–973.

ARROW, K. J. (1970) *Essays in the Theory of Risk-Bearing.* North-Holland Publishing Co., Amsterdam.

ARROW, K. J. (1974) Optimal insurance and generalized deductibles. *Scand. Actuar. J.* 1–42.

BEIGLBÖCK, M., GOLDSTERN, M., MARESCH, G., and SCHACHERMAYER, W. (2008) Optimal and better transport plans. Mimeo.

BORCH, K. (1962) Equilibrium in a reinsurance market. *Econometrica* **30**, 424–444.

BRENIER, Y. (1991) Polar factorization and monotone rearrangement of vector-valued functions. *Comm. Pure Appl. Math.* **44**, 375–417.

BRÉZIS, H. (1973) *Opérateurs Maximaux Monotones et Semi-groupes de Contractions dans les Espaces de Hilbert.* North-Holland Publishing Co., Amsterdam.

CAMBANIS, S., SIMONS, G., and STOUT, W. (1976) Inequalities for $Ek(X,Y)$ when the marginals are fixed. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **36**, 285–294.

CARLIER, G. and DANA, R.-A. (2005) Rearrangement inequalities in non-convex insurance models. *J. Math. Econom.* **41**, 483–503.

CHEW, S. and WAKKER, P. (1996) The Comonotonic Sure-Thing Principle. *J. Risk Uncert.* **12**, 5–27.

CUESTA-ALBERTOS, J. A., RÜSCHENDORF, L., and TUERO-DÍAZ, A. (1993) Optimal coupling of multivariate distributions and stochastic processes. *J. Multivariate Anal.* **46**, 335–361.

DALL'AGLIO, G. (1956) Sugli estremi dei momenti delle funzioni di ripartizione doppia. *Ann. Scuola Norm. Sup. Pisa* **10**, 35–74.

DENNEBERG, D. (1994) *Non-Additive Measure and Integral.* Kluwer, Dordrecht.

DHAENE, J., DENUIT, M., GOOVAERTS, M. J., KAAS, R., and VYNCKE, D. (2002a) The concept of comonotonicity in actuarial science and finance: applications. *Insurance Math. Econom.* **31**, 133–161.

DHAENE, J., DENUIT, M., GOOVAERTS, M. J., KAAS, R., and VYNCKE, D. (2002b) The concept of comonotonicity in actuarial science and finance: theory. *Insurance Math. Econom.* **31**, 3–33.

FRÉCHET, M. (1951) Sur les tableaux de corrélation dont les marges sont données. *Ann. Univ. Lyon. Sect. A. (3)* **14**, 53–77.

GALICHON, A. and HENRY, M. (2008a) Comonotonic measures of multivariate risks. URL `http://ssrn.com/abstract=1115729`.

GALICHON, A. and HENRY, M. (2008b) Dual theory of choice under multivariate risks. URL `http://ssrn.com/abstract=1126483`.

GANGBO, W. and MCCANN, R. J. (1996) The geometry of optimal transportation. *Acta Math.* **177**, 113–161.

GENEST, C., QUESADA MOLINA, J. J., and RODRÍGUEZ LALLENA, J. A. (1995) De l'impossibilité de construire des lois à marges multidimensionnelles données à partir de copules. *C. R. Acad. Sci. Paris Sér. I Math.* **320**, 723–726.

HARDY, G. H., LITTLEWOOD, J. E., and PÓLYA, G. (1952) *Inequalities.* Cambridge, at the University Press. 2d ed.

JOUINI, E. and NAPP, C. (2003) Comonotonic processes. *Insurance Math. Econom.* **32**, 255–265.

JOUINI, E. and NAPP, C. (2004) Conditional comonotonicity. *Decis. Econ. Finance* **27**, 153–166.

KAST, R. and LAPIED, A. (2003) Comonotonic book making and attitudes to uncertainty. *Math. Social Sci.* **46**, 1–7.

LANDSBERGER, M. and MEILIJSON, I. (1994) Co-monotone allocations, Bickel-Lehmann dispersion and the Arrow-Pratt measure of risk aversion. *Ann. Oper. Res.* **52**, 97–106.

LORENTZ, G. G. (1953) An inequality for rearrangements. *Amer. Math. Monthly* **60**, 176–179.

LUDKOVSKI, M. and RÜSCHENDORF, L. (2008) On comonotonicity of Pareto optimal risk sharing. *Statist. Probab. Lett.* **78**, 1181–1188.

MCCANN, R. J. (1995) Existence and uniqueness of monotone measure-preserving maps. *Duke Math. J.* **80**, 309–323.

MCNEIL, A. J., FREY, R., and EMBRECHTS, P. (2005) *Quantitative Risk Management: Concepts, Techiniques, Tools.* Princeton University Press, Princeton, NJ.

MEILIJSON, I. and NÁDAS, A. (1979) Convex majorization with an application to the length of critical paths. *J. Appl. Probab.* **16**, 671–677.

MINTY, G. J. (1962) Monotone (nonlinear) operators in Hilbert space. *Duke Math. J.* **29**, 341–346.

NELSEN, R. B. (2006) *An Introduction to Copulas.* Springer Series in Statistics. Springer, New York, second edition.

RACHEV, S. T. (1991) *Probability Metrics and the Stability of Stochastic Models.* John Wiley & Sons Ltd., Chichester.

RACHEV, S. T. and RÜSCHENDORF, L. (1998) *Mass Transportation Problems. Vol. I and II.* Springer-Verlag, New York.

ROCKAFELLAR, R. T. (1966) Characterization of the subdifferentials of convex functions. *Pacific J. Math.* **17**, 497–510.

ROCKAFELLAR, R. T. (1970a) *Convex Analysis.* Princeton Mathematical Series, No. 28. Princeton University Press, Princeton, N.J.

ROCKAFELLAR, R. T. (1970b) On the maximal monotonicity of subdifferential mappings. *Pacific J. Math.* **33**, 209–216.

RÜSCHENDORF, L. (1980) Inequalities for the expectation of Δ-monotone functions. *Z. Wahrsch. Verw. Gebiete* **54**, 341–349.

RÜSCHENDORF, L. (1991a) Bounds for distributions with multivariate marginals. In

Stochastic Orders and Decision under Risk (Hamburg, 1989), volume 19 of *IMS Lecture Notes Monogr. Ser.*, 285–310. Inst. Math. Statist., Hayward, CA.

RÜSCHENDORF, L. (1991b) Fréchet-bounds and their applications. In *Advances in Probability Distributions with Given Marginals (Rome, 1990)*, volume 67 of *Math. Appl.*, 151–187. Kluwer Acad. Publ., Dordrecht.

RÜSCHENDORF, L. (1996) On c-optimal random variables. *Statist. Probab. Lett.* **27**, 267–270.

RÜSCHENDORF, L. (2004) Comparison of multivariate risks and positive dependence. *J. Appl. Probab.* **41**, 391–406.

RÜSCHENDORF, L. and RACHEV, S. T. (1990) A characterization of random variables with minimum L^2-distance. *J. Multivariate Anal.* **32**, 48–54.

SCARSINI, M. (1989) Copulae of probability measures on product spaces. *J. Multivariate Anal.* **31**, 201–219.

SCHACHERMAYER, W. and TEICHMANN, J. (2008) Characterization of optimal transport plans for the Monge-Kantorovich-problem. *Proc. Amer. Math. Soc.* forthcoming.

SCHMEIDLER, D. (1989) Subjective probability and expected utility without additivity. *Econometrica* **57**, 571–587.

SCHWEIZER, B. and SKLAR, A. (1983) *Probabilistic Metric Spaces.* North-Holland Series in Probability and Applied Mathematics. North-Holland Publishing Co., New York.

SKLAR, M. (1959) Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris* **8**, 229–231.

TCHEN, A. H. (1980) Inequalities for distributions with given marginals. *Ann. Probab.* **8**, 814–827.

TOPKIS, D. M. (1998) *Supermodularity and Complementarity.* Frontiers of Economic Research. Princeton University Press, Princeton, NJ.

WAKKER, P. (1996) The sure-thing principle and the comonotonic sure-thing principle: an axiomatic analysis. *J. Math. Econom.* **25**, 213–227.

YAARI, M. E. (1987) The dual theory of choice under risk. *Econometrica* **55**, 95–115.

ZARANTONELLO, E. H. (1960) Solving functional equations by contractive averaging. Technical Report 160, Mathematics Research Centre, University of Wisconsin, Madison.

ZARANTONELLO, E. H. (1967) The closure of the numerical range contains the spectrum. *Pacific J. Math.* **22**, 575–595.

ZOLOTAREV, V. M. (1983) Probability metrics. *Theory Probab. Appl.* **28**, 278–302.

March 15, 2009